DOCUMENT RESUME

ED 191 563	PS 011 657
TITLE	Accept My Profile: Perspectives for Head Start Profiles of Program Effects on Children. Technical Report. Volume I (Revised Edition).
INSTITUTION	Mediax Associates, Inc., Westport, Conn.
SPONS AGENCY	Administration for Children, Youth, and Families (DHEW), Washington, D.C.
PUB DATE	Jun 80
CONTRACT	105-77-1006
NOTE	167p.: For Volumes II and III, see PS 011 658-659.
EDRS PRICE	MF01/PC07 Plus Postage.
DESCRIPTORS	Classification: Data Collection: *Developmental Tasks: Individual Characteristics: *Measures (Individuals): *Outcomes of Education: *Parent Attitudes: Parent Participation: *Preschool Children: Preschool Education: Profiles: Program Evaluation: Reliability: Research Methodology: Research Problems: *Teacher Attitudes: Validity
IDENTIFIERS	*Project Head Start: Test Batteries

ABSTRACT

This technical report discusses the development of a framework for identifying a battery of measures to be used in evaluating Head Start program effects on children 3 to 7 years of age. The report consists of an overview and five chapters, and provides eight appendices (in two additional volumes). Among the topics addressed are background information related to project planning and implementation: methodological procedures followed in determining characteristics to be measured: and issues affecting the selection of characteristics and the problems of defining constructs. Further, the specific characteristics recommended for measurement and the proposed system for their classification, as well as technical issues confronting test developers are discussed. The final chapter provides a summary of professional standards for reliability and validity, a discussion of other psychometric issues, a review of advantages and problems of using profiles and a discussion of several issues related to data collection procedures. (Author/RH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

TECHNICAL REPORT

VOLUME I

accept my profile

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Herman P.
Taub
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

perspectives for Head Start profiles of program effects on children

Prepared for the Administration for Children, Youth and Families, Office of Human Development Services, Department of Health, Education and Welfare.

Contract No: HEW-105-77-1006 Mediax Associates, Inc. Westport, Conn.
March, 1980
Revised: June, 1980

PROJECT PERSONNEL

Esther Kresh, Project Officer,
Administration for Children, Youth and Families

Staff

Herman P. Taub, Project Director
Louise Arcudi, Project Administrator
J. McVicker Hunt, Senior Scientist
Evelyn Keisman, Research Associate
Michael Linacre, Data Processing Specialist
John Love, Deputy Director
Ernest D. Washington, Child Development Specialist
Doxey A. Wilkerson, Technical Coordinator
James M. Wolf, Child Development Specialist

National Panel

Saundra Aldridge
Eli Bower
Jean Carew
William E. Cross
Edward A. DeAvila
Rosalie Dixon
Dorothy Eichorn
Roger Givens
William L. Goodwin
Susan Gordon
Asa G. Hilliard
Luis Laosa
Lee C. Lee
Ruth Marshall
Winona Sample
Jean Collins Sanford
Milton N. Silva

ACKNOWLEDGMENTS

The staff of this project is deeply grateful for the contributions of many people to its work. Especially notable among them are:

- Esther Kresh, ACYF project officer, who, more than any other person, is responsible for expertly shepherding this project from its birth through many crises, to its present stage.

- Many other members of the national administration and staff of the Administration for Children, Youth and Families, who have been consistently supportive, sharing their time and expertise unreservedly. Appreciation is expressed particularly to John Calhoun, Commissioner; Laura Miller, Deputy Commissioner; Barbara Kamara, Associate Commissioner; Blandena Cardenas Ramirez, former Commissioner; Sol Rosoff, former Deputy Commissioner; James Robinson, former National Head Start Director; Preston Bruce; Ray Collins; Dennis Deloria; and Clennie Murphy.

- Nearly 400 Head Start parents and staff members and K-2 teachers, who spent two days of intensive work in helping to identify developmental objectives for the children served by Head Start.

- Administrators and staffs of the several ACYF regions, who made it possible for the project to work with these Head Start parents and staff members and K-2 teachers.

- More than 100 respondents to requests for nominations for members of the National Panel.

- J. McVicker Hunt, whose continuing service as Senior Scientist makes available to this project, in an advisory capacity, his unparalleled expertise and insight in the field of early childhood development.

- Over 40 other leading specialists in the field of child development, who reacted to staff inquiries and proposals in individual conferences and through memoranda.

- The officers and directors of the National Head Start Associations, whose encouragement, advice, sharing of experiences and total dedication to children has served as a constant reminder of the important purpose of this project.

- The seventeen members of the National Panel, each of whom has committed a great deal of personal time and made valuable contributions to the tasks here addressed.

Whatever merit characterizes the ideas and proposals here advanced stems in large measure from the assistance of these and other participants in the project. Responsibility for the final document, however, lies only with Mediax Associates and, specifically, its director.

Herman P. Taub

CONTENTS

VOLUME I: TECHNICAL REPORT

Project Personnel i
Acknowledgements ii
List of Tables vii
List of Figures ix

OVERVIEW 1

The Project 1
This Report 1
Features of Proposed Measures 3
Characteristics and Structure for Measurement 7

I. INTRODUCTION 11

Background 11
    Origin of Head Start 11
    Head Start Goals and Performance Standards 12
    History of Evaluation Efforts 13
    Specific Antecedents of this Project 14

The Current Project 15
    Purpose 16
    Objectives 17
    Basic Premises 18
    Uses of Project Outcomes 19
    Non-Staff Personnel 20

Overview of Project Phases 20

II. METHODOLOGY 22

General Operating Procedures 22
    Staff 22
    National Panel 22
    Consultants 23
    Dissemination 25
    Relations with ACYF 26
    Summary 26

Defining the Issues 26

Determining Child Characteristics To Be Measured 28
    Scholar's Survey 28
    Head Start Input Workshops 28
    Literature Search 32
    Analysis of Findings 32

Analysis of Available Measures 33

III. CONSIDERATIONS IN DEFINING CHARACTERISTICS 36

Need for a Comprehensive Definition of Program Objectives 36
Need for Unbiased Indicators 39
Diversity 40
Problem of Defining Domains 41
Problems in Classifying Child Development Characteristics 42
Criterion, Background and Other Variables 44

IV. CHARACTERISTICS TO BE MEASURED 45

Approach to Defining Characteristics 45

Classifying Individual Characteristics 46
  Process-Content Categories 47

  Type Categories 51
    Knowledges and Skills 51
    Attitudes and Habits 52
    Physiological Statuses 52
    Executive Skills 53

Defining Domain Categories 53
  Cognitive Development 57
    Perception 59
    Language 61
    Reading 61
    Math Concepts 61
    Social Organization 61

  Social-Emotional Development 66
    Sensitivity to Feelings of Others 73
    Expression of Own Feelings 73
    Self-Concept 73
    Attitudes Toward School 76
    Independence 78
    Sharing and Competing 78
    Peer Relationships 78
    Adult Relationships 82

  Health and Physical Development 82
    Health and Safety 85
    Dental 87
    Nutrition 92
    Gross Motor Development 93
    Fine Motor Skills 95

  Applied Strategies 98

Rating of Characteristics 99
    Ratings of Importance 99
    Most Positively Rated Characteristics 103
    Most Negatively Rated Characteristics 103
    Characteristics with Neutral Ratings 107
    Demographic Group Differences in Importance Ratings 107

Age Ratings 112

V. TECHNICAL ISSUES FOR MEASURES DEVELOPMENT 115

Shortcomings of Existing Measures 115
    Criteria for Judging Measures 115
        Relevance of Content 116
        Psychometric Characteristics 116
        Practical Considerations 116
    Judgments of Members of the National Panel 117

    Conclusion 117

New Psychometric Challenges 118
    Challenges to Measurement Validity 118
        Accommodating Local Program and Diversity of Cultural/Ethnic Groups 118
        Accommodating Language Differences 119
        Accommodating Diversity in Patterns of Development 120
        Assessing Behavior in a Variety of Settings 121

    Applications of Standards of Validity and Reliability 121
        Validity Standards 122
        Reliability Standards 126
        The Application of Generalizability Theory 129

Issues in Data Analysis 130

    Item Analysis 130
    Profile Analysis 132
    Reporting Program Effects Without Reference to Norms 137
    Measurement Bias and Fairness 137
    Levels of Aggregation 138
    Units of Analysis 130
    Scaling levels of Development 140

Data Collection and Measurement Procedures 141
    Multiple Measurement Techniques 141
    Conditions for Eliciting Behavior 143
    Data Collection Issues 144

VI. PERSPECTIVES AND REASONABLE EXPECTATIONS 145

REFERENCES 146

VOLUME II: APPENDICES A - D (Bound Separately)

Appendix A: A System for Classifying Characteristics Identified In Head Start Input Workshops

Appendix B: Reduced List of Characteristics by Process-Content-Dimension (With Input Workshop References)

APPENDIX C: Reduced List of Characteristics by Content-Process-Dimension (With Input Workshop References)

Appendix D: Reduced List of Characteristics by Dimension-Process-Content (With Input Workshop References)

VOLUME III: APPENDICES E - H (Bound Separately)

Appendix E: Listing of Original Workshop Characteristics By Process-Content-Dimension (with Input Workshop References)

Appendix F: Tabulation of Importance Ratings and Differentiated Ratings by Demographic Group

Appendix G: Summary List of Process-Content Categories of Characteristics Identified with Dimension Constructs with Age Range Analysis

Appendix H: Comments of J. McVicker Hunt on the Cognitive Domain

LIST OF TABLES

Table No.		Page No.
I	Domains and Dimensions of Child Development Recommended for Measurement	10
II	Number of Participants in Input Workshops by Demographic Group	30
III	Facets of Social Competency	37
IV	Outcomes Associated with Social Competence	38
V	Number and Percent of Characteristics of Each Process Type	48
VI	Number and Percent of Characteristics of Each Content Type	49
VII	Attributes Recommended for Check-up and Treatment Through Medical Examination in the Health and Safety Dimension	88
VIII	50 Characteristics Rated Most Positive by Participants in Input Workshops	104
IX	Comparison of Importance Ratings of Specific Content Characteristics	105
X	50 Characteristics Rated Most Negative by Participants in Input-Workshops	107
XI	50 Chacteristics Rated Most Neutral by Participants in Input-Workshops	108
XII	Summary of Positively and Negatively Differentiated Importance Ratings of Demographic Groups	110
XIII	Summary of Characteristics Consistently (75%) Rated Differentially from Ratings of Overall Participants	111

XIV	Distribution of Mean Age Ratings by Input-Workshop Participants of Characteristics Hypothesized as Associated with Each Dimension	113
XV	Reliability and Validity Standards for Educational and Psychological Tests	123
XVI	Possible Sources of Variance of Scores on a Particular Test	127

LIST OF FIGURES

Figure No.		Page No.
1	Conceptualized Relationship of Desired Developmental Characteristics and Domains and Dimensions of Comptency	8
2	Process and Content of Characteristics Identified in All Domains	50
3	Types of Developmental Characteristics to be Measured in the Four Domains	56
4	Process and Content of Characteristics Identified in the Cognitive Domain	58
5	Process and Content of Characteristics Identified in Dimension 51: Perception	60
6	Process and Content of Characteristics Identified in Dimension 52: Language	62
7	Process and Content of Characteristics Identified in Dimension 53: Reading	63
8	Process and Content of Characteristics Identified in Dimension 54: Math Concepts	64
9	Process and Content of Characteristics Identified in Dimension 55: Nature and Science	65
10	Process and Content of Characteristics Identified in Dimension 56: Social Organization	67
11	Process and Content of Characteristics Identified in the Social-Emotional Domain	71
12	Process and Content of Characteristics Identified in Dimension 61: Sensitivity to the Feelings of Others	74

Figure No. Page No.

13 Process and Content of Characteristics Identified in Dimension 62: Expression of Own Feelings 75

14 Process and Content of Characteristics Identified in Dimension 63: Self-Concept 77

15 Process and Content of Characteristics Identified in Dimension 64: Attitudes Toward School 79

16 Process and Content of Characteristics Identified in Dimension 65: Independence 80

17 Process and Content of Characteristics Identified in Dimension 66: Sharing & Competing 81

18 Process and Content of Characteristics Identified in Dimension 67: Peer Relationships 83

19 Process and Content of Characteristics Identified in Dimension 68: Adult Relationships 84

20 Process and Content of Characteristics Identified in Health & Physical Development 86

21 Process and Content of Characteristics Identified in Dimension 71: Health and Safety 89

22 Process and Content of Characteristics Identified in Dimension 72: Dental 90

23 Process and Content of Characteristics Identified in Dimension 73: Nutritional 94

24 Process and Content of Characteristics Identified in Dimension 74: Gross Motor 96

Figure No.		Page No.
25	Process and Content of Characteristics Identified in Dimension 75: Fine Motor	97
26	Process and Content of Characteristics Identified in the Applied Strategies Domain	100
27	Process and Content of Characteristics Identified in Dimension 81: Task Competencies	101
28	Process and Content of Characteristics Identified in Dimension 82: Interpersonal Competencies	102
29	Profile of Overall Program Effects: Sample Column Diagram	133
30	Profile of Overall Program Effects: Sample Sector Graph	134
31	Profile of Program Effects in a Single Dimension: Sample Column Diagram	135
32	Profile of Program Effects in a Single Dimension: Sample Sector Graph	136

OVERVIEW

This report is the culminating effort of Phase I of a project to develop Head Start profiles of program effects on children. The specific objective of the overall project is to develop a battery of measures suitable for evaluating the effects the Head Start Program on children between the ages of 3 and 7. It is not the purpose of this project to develop measures that are suitable for the assessment of individual children.

The Project

The project has been planned to be carried out in three phases.

In Phase I the characteristics to be measured have been identified, a structure for an integrated system of measurement was defined, and the need for new measures development was assessed. Also, the Administration for Children Youth and Families (ACYF) was presented with various options for undertaking the development of a recommended measurement system. After ACYF selected a course of procedures, specifications for the procurement of measures development services, as required, were prepared. During Phase II, the recommended measures will be prepared, and the development and pilot testing of the new system of measures is to be completed. In Phase III a national field test of the overall measurement system is planned to be carried out with Head Start programs.

Phase I, which started in September 1977, is scheduled for completion in November 1980. Phase II is planned to run from November 1980 through 1983. Phase III is to start prior to the beginning of the Head Start program year in 1983.

This Report

This technical report is organized in an overview and five chapters included in this volume, and eight appendices (A through H) in two additional volumes of related appendices.

This overview includes summaries of the distinctive characteristics of the proposed measures (pp. 3-6) and of the proposed domains, dimensions and structure of the measurement system (pp. 7-10).

Chapter I, Introduction provides background information to this report. It describes those forces and events which led to the initiation of Head Start and this project. It attempts to identify the specific programmatic, administrative and political issues affecting this project. It describes the purposes, objectives and basic premises of the project. Particular emphasis is placed on differentiating appropriate uses of the outcome measures for the evaluation of overall program effectiveness and inappropriate uses that involve assessment of individual children. Chapter I also includes a description of the specific tasks for each phase of the project.

Chapter II, Methodology, describes the procedures followed in carrying out the tasks leading to this report. Of particular note is the formation and membership of a National Panel (pp. 22-23) to serve in an advisory capacity and to oversee each of the project's activities.

Chapter II also describes the commitment of the project to broad dissemination of its procedures and findings and to the participation of outstanding child development specialists representing divergent points of view; the specific consultants who have been involved in the fulfillment of that commitment are listed (p. 24).

Three approaches followed in determining characteristics to be measured are documented in Chapter II: (1) the survey of 22 scholars (p. 28); (2) the input workshops involving 375 Head Start parents, staff and public school teachers (pp. 28-32); and (3) a literature search. Finally, the documents and procedures used in the analysis of available measures are described (pp. 33-35).

Chapter III, Considerations in Defining Characteristics, outlines the most important issues affecting the selection of characteristics. The findings of particular antecedent projects are summarized, and the need for more explicit statements of indicators of desired Head Start effects on children is explained; of special concern are indicators that represent the judgments of persons familiar with the culture and developmental paths of particular children and that encompass the diversity in goals and values of the various populations in Head Start.

Also in Chapter III are discussions of problems in defining constructs; the need for a system to classify characteristics and the problems in establishing one; and the need for criterion, background and other variables.

Chapter IV, Characteristics to be Measured, describes and analyzes the specific characteristics recommended for measurement and the proposed system for their classification. All characteristics are classified by their process and content properties and by their type: knowledges and skills, attitudes and habits, physiological statuses or executive skills. These various categorizations are used to describe and analyze each of the domains and dimensions recommended for measurement; they are also summarized on pages 7 through 10 of this overview.

In Chapter V, Technical Issues for Measures Development, some of the technical issues confronting measures developers are discussed; particularly those special requirements imposed by the need to accommodate local program emphases, language differences, and diversity in children's development. Here, also, the procedures and findings of the review carried out in this project of existing measures are further discussed. The criteria are identified by which available measures were found inadequate and by which new measures will be judged, and the need for a more comprehensive and coherent system is explicated.

Also in Chapter V is a summary of professional standards for reliability and validity; the application of these and other standards to the measures to be developed is discussed. Psychometric issues such as generalizability theory, item analysis, requirements for norms, aggregation techniques and the need for scaling techniques that reflect the diverse paths of development of Head Start children are also discussed.

The advantages and problems of using profiles is reviewed; alternative conceptualizations of graphic profiles are included. Finally, there is a discussion of the need for multiple measurement techniques, of the effects of the conditions used to elicit behavior for measurement, two important data collection issues: the time required for administration, and the recommendation that measures be administered by paraprofessionals. A maximum of one hour (divided into two half-hour sessions) is recommended for directly involving any individual child at the Head Start level.

This paper ends with a brief statement on perspectives and reasonable expectations.

Features of Proposed Measures

The various investigations, analyses and discussions undertaken in this project will result in a system of proposed measures which differ from those existing in the following major respects:

1. They will encompass all Head Start objectives.

   To date, evaluators of Head Start have had to piece together measures from a variety of sources in an attempt to measure all the Head Start objectives for children. These measures were developed independently, and were often designed for other purposes and based on competing theories and philosophies of child development.

   Early attempts at measuring Head Start effects made use of intelligence tests which give considerable value to children's past experiences in a white, middle-class culture. This led some observers to erroneous and dangerous conclusions about the developmental potential of children of the poor and about poor minority children in particular.

   A second source of measures was developmental screening tests. These instruments are designed for identification of possible developmental abnormalities or lags in children's normal course of development. While useful for their intended purposes, they are inadequate measures of program effects as they are limited to characteristics representative of normal but not necessarily teachable behaviors. Therefore, with the use of such measures on the large majority of Head Start children, no differences due to program effects would be expected.

   A third set of measures that has been used can be classified as general achievement measures. These have had the most utility, but they suffer from limited scope and questionable applicability to all of the Head Start population. Such existing cognitive measures, for example, are designed to report the attainment of a narrow band of desired characteristics rather than being sensitive to the wide range and levels of skills attained by children entering Head Start.

   In the area of social-emotional development, evaluators have consistently been frustrated in finding existing instruments which are reliable and valid for the Head Start population. This has led to a disproportionate emphasis on cognitive skills that are more readily measurable, leading many people, not entirely knowledgeable about Head Start, to assume that the program's primary interest is in enhancing cognitive development as opposed to Head Start's more comprehensive objectives.

2. They will be tailored to the Head Start objectives as set forth in the "Performance Standards" and the specific inputs from Head Start parents and staff.

On a general level, Head Start objectives are as set forth in the Head Start Performance Standards, which were derived from the philosophical position of the architects of the program and the persons who administer it. On a more specific level, Head Start is explicitly committed to local articulation of children's behaviors and status characteristics to define each of the broad objectives. It is inherent in the process of children's development that parents and local community, through their everyday interactions with their children, influence the characteristics which are indicative of desired development.

This project, therefore, involved Head Start parents, staff and public school teachers (of Head Start graduates), representative of all the various ethnic backgrounds, regions and life styles[1] served by Head Start, in identifying these characteristics. Since most available measures were developed independent of concerns of parents, staff and teachers of Head Start children, the measures here developed will be the first to purposefully address characteristics valued by the people most affected by the Head Start program and, therefore, representative of desired development of Head Start children.

However, because these measures are being developed specifically for Head Start children, they should not be assumed to be applicable to the measurement of effects of other programs. A similar process to that undertaken in this project would be required to either validate or redefine the domains of desired development in terms of characteristics appropriate for the particular populations served by such other programs.

3. The measures will allow for diversity in child development.

Although a number of existing measures now have norms on different racial/ethnic groups, none that have been used in previous Head Start evaluations were actually developed to insure that culturally different expressions of similar outcomes were included. The required measures must be based on valid indicators for desired outcomes that are appropriate, and may be different, for the various racial/ethnic/cultural/regional groups in Head Start programs. For example, "looking the teacher in the eye" may be a desired expression of attention in some cultural groups; whereas attention may be expressed in completely different ways in another.

---

[1] The term "life styles" is used in this context to refer to the broad environmental context in which people live, e.g., rural or urban settings, migrant life patterns, etc.

All measures-development efforts in this project will, therefore, use data from the three main Head Start populations (Black, Anglo, and Hispanic) and at least one other ethnic/cultural group, and various regional subgroups thereof, to insure that all appropriate indicators representative of selected outcomes are included. In addition, data analysis procedures will be developed to assure valid interpretations for particular groups, where appropriate. The requirement that each measure must utilize Black, Anglo and Hispanic groups in its development stems from the fact that over 94 percent of the Head Start population consists of these three groups (National Institute for Advanced Studies, 1980). In the opinion of the members of the National Panel, if measures are appropriate for the three main groups and one additional group, they will more than likely be sufficiently appropriate to be used in evaluations that encompass all other groups in Head Start. Due to limited resources, where a measure involves language, only a Spanish language alternative is being recommended at this time. These techniques will result in greater appreciation for and, thereby, enhancement of the diverse patterns of desired development among Head Start children.

4. The measures will express program effects in terms of profiles of change.

Available measures produce scores which, most often, represent quantitative aggregations of data on different characteristics, thereby concealing important differences. Also, such scores are often calibrated to different scales. A change in score of 30 points in one measure might represent more, less, or indeterminable growth relative to a change in score of 130 points on another. These limitations lead to the often erroneous conclusion that "more of everything is better".

The measures of this project will be designed to communicate program effects in terms of profiles of changes in each of the areas of child development comprising overall desired outcomes. They may be expected to be sensitive to qualities and patterns of competency previously undetected.

5. To the extent possible, the measures will use scales that are developmentally sequential or hierarchical.

Most of the measures that have been used in Head Start evaluations to date are not designed so that a given score is translatable to the attainment of specific characteristics within a domain. That is, the scores children receive do not give any information about what those children can or cannot do. For example, given a measure based on 50 items, two scores of 25 could represent mastery of entirely different sets of items. Although the scores are the same, they may represent knowing or doing different things.

Although not every area to be measured has a sequential or hierarchical pattern of development, wherever possible, an attempt will be made to develop scales on which attainment of any given scale point will indicate that a child could do all the things below that point and none of the things above. It is unlikely that the state-of-the-art will permit complete success in developing such scales, but there is good reason to believe that an approximation of such scales can be made.

6. Different techniques will be employed to measure the same characteristics.

In many of the Head Start evaluations, an attempt has been made to employ simultaneously different techniques, e.g., direct observations of children with parent and/or teacher ratings. However, such techniques have most often come from different sources (e.g., an already developed observation system combined with a contractor-developed parent and teacher rating scale) and, therefore, they usually include information on different sets of characteristics.

The new measures will include different measurement techniques developed for the same set of characteristics, where appropriate. The result should be a greater level of precision in the measurement of child outcomes through the reduction of such distortions as measurement error and bias.

7. Wherever appropriate, the measures will be based on characteristics as they are manifested in different settings, such as the home, school and community.

It is obvious that for some characteristics the specific indicators are the same in all settings (e.g., identifying colors, frequency of sickness, etc.). Other characteristics may occur only in a particular setting (e.g., eating or sleeping habits), or manifest themselves differently from setting to setting (e.g., relationships with adults). Most existing measures are designed to be administered in only one setting and, therefore, would be insensitive to program effects on characteristics manifested in other settings.

The measures here will expand the settings in which they can be applied, thus expanding the behavioral repertoire of children that can be measured.

8. The measures will be developed for use with children aged 3-7.

At present, most existing measures that span the entire age range of interest (3-7) are developmental screening tests. The limitations of such measures were discussed earlier in this paper. Other available measures are appropriate only for some preschool or school years. Measures that span continuously the years from 3 to 7 must encompass the range of characteristics representative of a period of particularly rapid development and must also cover developmental transitions, as from the preoperational to the concrete operational stages.

By developing a system of measures that provides continuous coverage of a wide range of characteristics over these years, it is possible to establish a data-base on program effects that can lead to important insights on longitudinal relationships.

Characteristics and Structure for Measurement

Through the various procedures described in the following chapters of this report a total of 1,516 characteristics are recommended for consideration for measurement. These define a structure of domains and sub-domains which are the basis for analyzing and reporting overall program effects.

The structure developed in this project evolved through an analysis of the characteristics identified in the Head Start Performance Standards, a survey of child development scholars, the recommendations from two earlier ACYF studies (Anderson & Messick, 1973; Raizen & Bobrow, 1974), and the specific outcomes identified in a survey of parents, staff, and public school teachers of Head Start children. Six broad categories, termed domains, appeared to cover all aspects of desired development. These domains have been labeled (1) the cognitive, (2) the social-emotional, (3) the health and physical, (4) the moral and ethical, (5) the aesthetic, and (6) the domain of applied strategies. The sixth domain, applied strategies, appeared to involve coordination of characteristics in all of the other domains and is, therefore, conceptualized as overarching the others.

With full appreciation for the importance of characteristics involving (4) moral and ethical development and (5) aesthetic development, these two domains were deleted from the concerns of this project because they are not priority objectives as defined in the Head Start Performance Standards at this time. Also, their measurement involves less studied techniques than are available for the other domains. Therefore, undertaking their measurement would divert significant amounts of limited resources from higher priority concerns.

The relationship between the remaining four domains and the types of characteristics conceptualized as existing under each is illustrated in Figure 1. These relationships and the overall structure will be reevaluated on the basis of data from the measures to be developed. For the present, the domains have been defined to include the following:

1. The Cognitive Domain

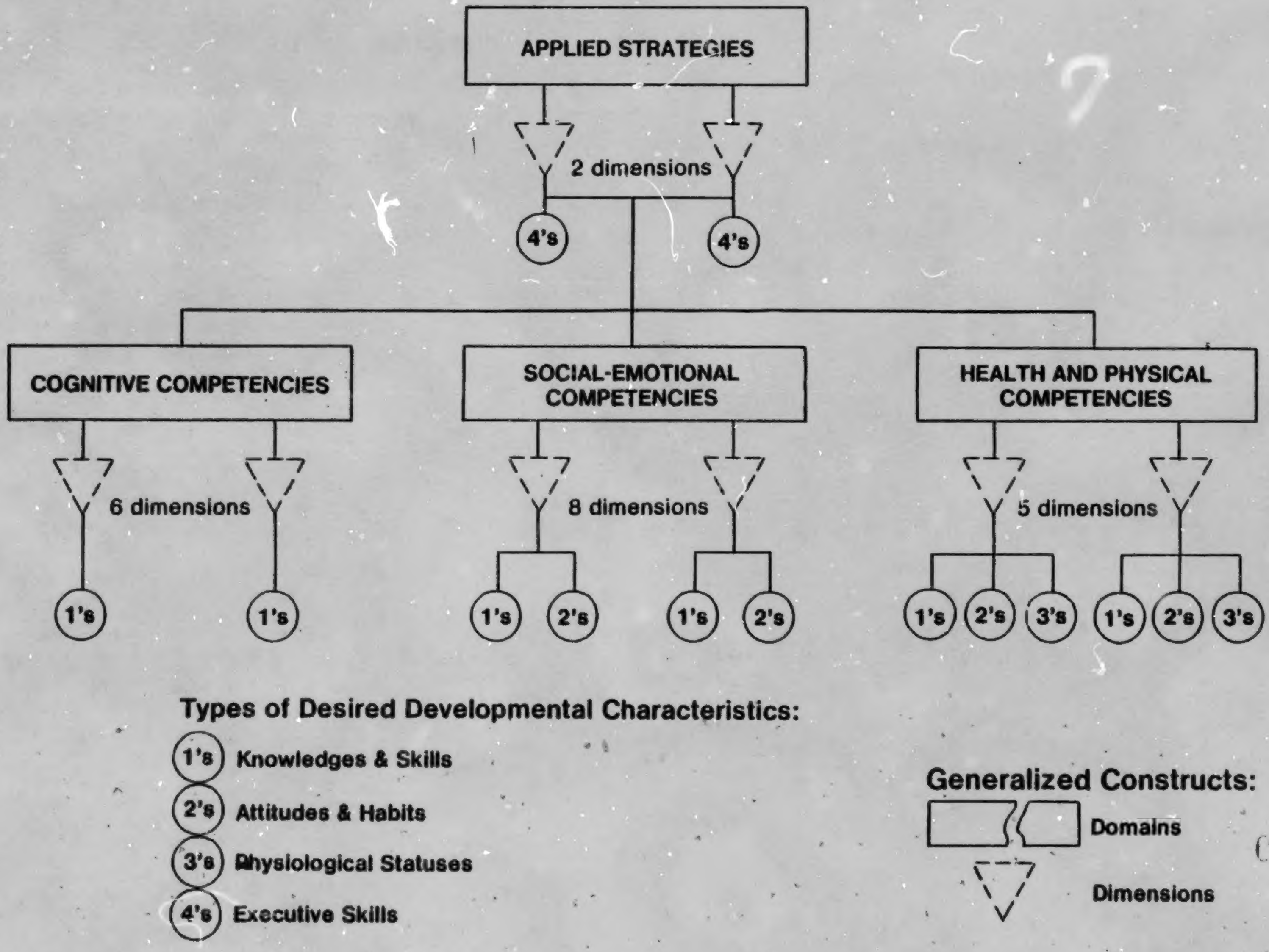
This domain includes language, both productive (speaking and writing) and receptive (reading and listening), mathematics, and knowledge of family, community, the workings of society, and the basic sciences. It includes both the acquisition of factual knowledge and the ability to apply it. The domain also includes the use of higher mental processes, such as comprehension and problem-solving.

2. The Social-Emotional Domain

Included in this domain are the characteristics that apply to the intrapersonal and interpersonal aspects of one's life. It includes an increasing capacity to know and be objective about oneself, awareness of feelings, attitudes, and the ability to participate in, adjust, and contribute to a given community.

Figure 1

Conceptualized Relationship of Desired Developmental Characteristics and Domains and Dimensions of Competency


APPLIED STRATEGIES
2 dimensions
4's
4's
COGNITIVE COMPETENCIES
SOCIAL-EMOTIONAL COMPETENCIES
HEALTH AND PHYSICAL COMPETENCIES
6 dimensions
8 dimensions
5 dimensions
1's
1's
1's
2's
1's
2's
1's
2's
3's
1's
2's
3's
Types of Desired Developmental Characteristics:
1's Knowledges & Skills
2's Attitudes & Habits
3's Physiological Statuses
4's Executive Skills
Generalized Constructs:
Domains
Dimensions

Note: Domains-Dimensions of Aesthetic and Ethical Development have been omitted from this figure.

3. The Health and Physical Domain

This domain includes those outcomes that contribute to one's physical well being. These constitute health, nutrition, small and gross motor development and increasing acuity of the sensory organs.

4. The Applied Strategies Domain

This domain is of a different order than the other three. It includes those behaviors that involve what we call executive skills in the integration and organization of the characteristics subsumed under the other domains, to devise and implement effective courses of action in specific real life situations. Of particular concern to Head Start are those attributes associated with effectiveness in learning situations, such as motivation, curiosity, initiative, persistence and task orientation, which are, therefore, instrumental to development in all of the other domains.

Under each of these broad domains, sub-domains (termed "dimensions") have been identified. The dimensions were selected so that each encompasses a somewhat unique set of characteristics in a commonly differentiated area of child development and that, together, encompass all of the characteristics elicited as indicators of desired development in the input workshops. The domain/dimension structure finally adopted is outlined in Table I.

This structure does not encompass anything that has not been previously identified as a desirable outcome for children. What is new is that this structure is the first based on large-scale inputs from parents and that it attempts to define all of the broad domains of desired characteristics in a single comprehensive system.

TABLE I

Domains and Dimensions of Child Development Recommended for Measurement

I. Cognitive Development

1. Perception
2. Language
3. Reading
4. Math Concepts
5. Nature and Science
6. Social Organization

II. Social-Emotional Development

7. Sensitivity to Feelings of Others
8. Expression of Own Feelings
9. Self-Concept
10. Attitudes Towards School
11. Independence
12. Sharing and Competing
13. Peer Relationships
14. Adult Relationships

III. Health and Physical Development

15. Health and Safety
16. Dental
17. Nutrition
18. Gross Motor
19. Fine Motor

IV. Applied Strategies

20. Task Competencies
21. Interpersonal Competencies

I

INTRODUCTION

Background

This Head Start Profiles of Program Effects on Children project was initiated in 1977 by the Administration for Children, Youth and Families in response to a need to develop more effective and appropriate tools for the evaluation of the Head Start Program and its related demonstrations. Although evaluations of Head Start have been undertaken almost from its inception, they have been hampered by a lack of measures which reflect the overall objectives of Head Start and are appropriate for the diverse cultural groups which Head Start serves.

The development of measures of program effects on children is seen as the first of a series of efforts to develop a comprehensive set of measures to assess all the areas Head Start attempts to affect -- children, parents, communities and institutions. This project is only a beginning. Ultimately, measures must also be developed for parents, communities and institutions. It is only then that a true assessment of the total impact of Head Start can be undertaken. The decision to start with measures of program effects on children was not derived from a higher priority being placed on these outcomes but rather because the area of child outcomes has been most studied -- and more often than not, with limited or inappropriate measures.

This project is the latest step in a development that has been under way for many years. Some of its important antecedents are discussed below.

Origin of Head Start

Head Start, launched in 1965 as part of the war on poverty, was conceived out of three major premises. The first consideration involved the ethical assumption of equal opportunity for all citizens in the United States. The second involved the recognition of the serious needs of children from families of poverty; i.e., it is the children of poverty who most commonly fare badly in school and drop out before they have achieved the credentials and skills required for employment in society. The third consideration was the evidence of plasticity in the psychological development of children.

Rejection of the viewpoint that psychological development was largely determined by hereditary factors in favor of the influence of environment was influenced by the work of two scholars, J. McVicker Hunt and Benjamin Bloom. "Hunt (1961) in his book Intelligence and Experience, abandoned the notion of fixed intelligence and abilities in favor of a new orientation emphasizing the power of a child's environment -- particularly the quality of mothering -- on intellectual growth" (Zigler & Valentine, 1979). In Stability and Change in Human Characteristics, Bloom (1964), marshalling longitudinal research conducted throughout the world, found that intellectual growth occurred most rapidly during the first four or five years of life and that variability in an individual's performance was found to be the greatest during the early years. He concluded that the best time to enrich the environment and affect intellectual growth was during this period of rapid development, i.e., the preschool years.

With these premises in mind, it was apparent that some form of intervention was urgently needed for children from low-income families if this nation were to fulfill its obligation of equal opportunity for all its citizens. Research studies indicated that a critical period in children's development encompassed the 3-to-5-year age range. The utilization of preschool experience was therefore considered to be the most efficient method of delivering services to young children who have been raised in poverty families.

Project Head Start was initiated as an eight-week, part-day program of experiences prior to school entry, which consisted of six components: early childhood education, parent involvement, health services, mental health services, nutrition and social services. This program was soon expanded to a full year. Also, an ongoing commitment was undertaken to improving overall Head Start effectiveness. This involved initiation of a series of pilot innovative alternatives at a limited number of sites, focusing, respectively, on various aspects of the interrelated effects of long-term child outcomes, improvement in family and community environments, and improvement in the institutions serving children, family and community. Each included an independent, thorough evaluation leading to possible recommendations for improvement of the Head Start program and for viable options for the delivery of services.

Head Start Goals and Performance Standards

The basic goals for Head Start have not changed since Congress initiated the program. In 1973, the first Head Start Policy Manual was issued, which outlined minimum performance standards for the achievement of these goals. Six goals are stated in the most recent edition of the Head Start Program Performance Standards (OCD, 1975b):

1. The improvement of the child's health and physical abilities, including appropriate steps to correct present physical and mental problems and to enhance every child's access to an adequate diet. The improvement of the family's attitudes toward future health care and physical abilities.

2. The encouragement of self-confidence, spontaneity, curiosity, and self-discipline, which will assist in the development of the child's social and emotional health.

3. The enhancement of the child's mental processes and skills with particular attention to conceptional and communications skills.

4. The establishment of patterns and expectations of success for the child, which will create a climate of confidence for present and future learning efforts and overall development.

5. An increase in the ability of the child and family to relate to each other and to others.

6. The enhancement of the sense of dignity and self-worth within the child and his family.

At this writing (June 1980), new performance standards are being reviewed within ACYF; no major changes are expected in the above stated goals.

History of Evaluation Efforts

The first overall evaluation of Head Start was mandated by Congress, funded by the Office of Economic Opportunity (OEO) and carried out by Westinghouse Learning Corporation and Ohio University, (Circirelli, Cooper & Granger, 1969). This report, utilizing children's standardized test scores as the outcome variables, identified modest short-term cognitive gains, but concluded that any advantage enjoyed by Head Start students had faded out by third grade. Although the report was criticized on technical grounds, it had important policy consequences.

One policy shift included new emphasis on experimentation and evaluation within Head Start. Head Start Planned Variation (HSPV) was initiated and included attempts to try out and compare alternative Head Start curricula during the preschool years. The Home Start demonstration was started in 1971 to obtain comparable educational and developmental gains for preschool children by working with mothers in their own homes, as are obtained by teachers working in Head Start Centers. Other innovative demonstration programs included Parent Child Centers (PCCs), Parent Child Development Centers (PCDCs), Project Developmental Continuity (PDC), Child and Family Resource Program (CFRP), services for handicapped children, the Child Development Associate (CDA) program, bilingual-bicultural programs for Spanish-background children, Head Start collaboration with Medicaid in effecting Early and Periodic Screening, Diagnosis, and Treatment (EPSDT) and, most recently (1979), improvement in children's Basic Educational Skills (BES).

Evaluation of the effects of each of these innovative programs and of overall Head Start was an important component of the new emphasis. The Educational Testing Service was contracted to do a longitudinal study of Head Start children from 4 to 8 years of age. The study compared Head Start children's cognitive, personal, and social development with that of children with other preschool experiences, and also with no preschool experiences, based on information collected on 1,212 children and their families (Shipman, Boroson, Bridgeman, Gant & Mikovsky, 1976). It concluded that children from poverty backgrounds appeared to develop more slowly than middle-class peers. However, this study was criticized for its design and conclusions in that "considerable difficulty is experienced in attributing outcomes to Head Start or other program treatment because of confounding with site, socio-economic status, and of those variables with race" (Collins, 1977, p. 363).

In 1977, Abt Associates carried out a nationwide study in 32 communities to evaluate the performance of former Head Start children and their peers who had not attended preschool. The study focused on the child's adaptation to and performance in the school environment, based on the rationale that ". . .program components of health, family, and social service may be seen ultimately as support systems for the child's adaptation to school and, thereby, later to the larger society" (Royster, et al 1977, p. 125). The study attempted to measure the following school behaviors: academic skills, motivational interest and persistence, pro-social adaptive behavior, unsocial regressive behavior, and dependency behaviors. Because of methodological problems, the findings of this study provided little evidence on what aspects of preschool experience account for the success of Head Start in improving school performance of children during the first year of school. However, the overall results did indicate that Head Start does lead to positive school achievement.

Attempts to measure program effects through techniques other than direct measurement of children have been more successful. The Administration for Children, Youth and Families funded research that resulted in the establishment of the Developmental Continuity Consortium Study Center with responsibility to analyze longitudinal data, collected over a ten year period, on children in twelve separate early-childhood education programs which operated in the early 1960's. Their analysis resulted in the identification of sustained positive effects of preschool programs on school readiness and achievement and, even more impressively, on referrals to special education classes and retention in grade (Lazar, Hubbell, Murray, Rosche & Royce, 1977).

In reviewing the past studies of Head Start and its related programs, a recurrent theme appears -- the evaluations have suffered from inadequately implemented research designs, inadequate and inappropriate measures, field problems in data collection, and too narrow a range of outcomes being assessed. All of these problems have contributed to the difficulty of confirming program effects. While each successive study has attempted to improve upon the methodology of those that went before, the problem of locating instruments which adequately measure the broad range of outcomes implied by Head Start goals and which also are maximally sensitive to detecting program effects has not been solved.

Evaluations have been frustrated by their inability to measure program effects with any confidence, other than short-term cognitive gains, despite the fact that there is general agreement that the most important long term effects are in such areas as social and emotional growth. Typical is this statement reviewing the Head Start Planned Variation evaluation: "The HSPV evaluators devoted considerable effort to a search for adequate non-cognitive measures. That they were unable to find them results not from a lack of will or effort, but from the lack of adequate theory and technology in the field" (Walker, Bane & Bryk, 1973).

Specific Antecedents of this Project

An early attempt to define the desired overall effects on children of Head Start was made by Edward Zigler, the first Director of the Office of Child Development. It was Zigler's position that Head Start should not be committed to a narrow cognitive development but should encourage a wider whole-child approach. Higher IQ scores and higher elementary school grades should not be the ultimate goal of the Head Start effort. "Let us proclaim instead that the primary goal of Head Start is to promote socially competent human beings" (Zigler 1977, p. 4). He elaborated on the meaning of 'social competence' as "an individual's everyday effectiveness in dealing with his environment. A child's social competence may be described as his ability to master appropriate formal concepts, to perform well in school, to stay out of trouble with the law and to relate well to adults and other children" (Zigler; Quoted in Anderson & Messick, 1973, p. 3). Zigler identified three specific components of social competence: a child's health, intellectual ability, and social and emotional development.

Thus, in 1972-1973, the construct "social competence" was introduced into Head Start. Social competence was defined by the Office of Child Development (now ACYF) as the overall goal of Head Start, to mean "the child's everyday effectiveness in dealing with his environment and later responsibilities in school and life. Social competence takes into account the interrelatedness of cognitive and intellectual development, physical and mental health, nutritional needs, and other factors that enable a child to function optimally" (OCD, 1975b, p. 1). Thus, social competence emerged as the critical dependent variable for program impact evaluation of Head Start.

The Educational Testing Service undertook to explicate the definition of the construct "social competence". "On January 23-25, 1973, a panel of experts met in Princeton, New Jersey to try to define the meaning of social competency in young children. They were united in feeling that social competency was something more than general intelligence and in recognizing the need to explicate that 'something more' as a basis for fostering and evaluating children's development" (Anderson & Messick, 1973, p. 1). As an outcome of this meeting, 29 "facets of social competence" were identified as encompassing all the domains of Head Start concern.

Subsequently, the Rand Corporation was asked to design a national evaluation of the social competence effects of the Head Start program on eligible children. The charge was to define "the major requirements for a national evaluation that uses social competence as the criterion variable, respects the rights of multiple constituents of the evaluation, and yields meaningful data" (Raizen & Bobrow, 1974, p. iii). In its final report, the project staff pointed out that the definition of social competence must be a function of the specific roles that are required of the child and must be measured in relation to the context in which those roles are manifested. Most Head Start children have to deal in two social contexts: in that of their own cultural group and in the society and institutions of the dominant culture. Rand saw Head Start, therefore, as a "preface" for children to their assuming the role of pupil as defined by the dominant culture. By this perception, Head Start's function was to prepare chldren for the transition from their own environment and culture to that of the school. Accordingly, the Rand report restricted the definition of social competence to "effectiveness in the role of pupil" (Raizen & Bobrow, 1974, pp. 17-18).

The specific characteristics identified in these ETS and Rand studies are listed in the next chapter on defining characteristics (pp. 37-38). These studies are referred to as "preliminary projects" in the overview to the specifications for this project. The specifications state, "Although OCD has continually maintained evaluation efforts of its early childhood programs, these evaluations have been hampered by a lack of adequate and appropriate instruments to measure the effectiveness of these programs on all areas of social competence. . .", therefore, the outcome desired from this project is". . .the compilation of a battery of social competence measures for Head Start children" (Department of Health, Education and Welfare, 1977).

The Current Project

The mission defined for this project -- "compilation of a battery of social competence measures for Head Start children" -- thus emerged from the limitations of previous efforts related to this end. Its goals, premises and approaches have been influenced greatly by this background.

Purpose

The utlimate purpose of the battery of social competence measures for Head Start children is the improvement of programs for children. It assumes the purpose of preschool and school programs in the United States is -- or should be -- to foster the development of human beings who are "at home" with themselves and their world, and who are capable of functioning effectively in the varied and changing circumstances that impinge upon their lives. The goal is what Head Start terms "social competence". It includes developing all the knowledges, skills, attitudes, habits, and physiological conditions that are basic to effective functioning in society.

This, likewise, is -- or should be -- the central focus of measurement and evaluation. Their ultimate value must be based on whether they yield insights that contribute to the development of generally competent human beings. Hence, the purpose of this measurement-development project is to provide "tools" which evaluators can use to assess the effects of Head Start on young children, which can then be used to identify practices which foster or frustrate desired development of children.

What is "desired development" thus becomes a key question in the measurement and evaluation of program effects; and it is answered differently by adherents to different theories and values. It is interpreted here in terms of the extent of growth toward objectives defined, qualitatively, by the Head Start parents and teaching staff. The focus is on the improvement in children's developmental status over the program period, i.e., the rate of development rather than the particular level attained. The responsibility for effecting a satisfactory overall rate of development, therefore, is placed on the fostering institution, where it belongs.

Although the growth of children is measured in this project, such measurement is used for assessment of program effects, not the appraisal of individual children. The effective evaluation of individuals requires that instruments and procedures exhibit validity and reliability for decisions involving individuals and, in that sense, are different than those for the evaluation of programs. Therefore, the new measures will require somewhat different psychometric properties than are required to appraise the status of individual children.

There are also important limitations in the purposes for which the measures of this project are being developed. While concerned with the improvement of programs, they are only intended for, and should be used on, Head Start or other programs with similar populations and objectives. The definitions of the domains and desired development used in this project can only be assumed to be valid for the populations for which they were developed. And the selection of indicators of this desired development is based on inputs of Head Start parents, staff, and public school teachers of Head Start children. Accordingly, measures of the effects of programs serving other populations should be based on definitions of desired development and program effects indicators that are derived from studies involving such other populations.

Also, it should be noted that the Head Start population includes a sizeable proportion of children (over 10 percent) who are classified as handicapped. Although many of the measures developed in this project may be suitable for many handicapped children, it is beyond the scope of this project to specifically address the problems of assessing program effects on each category of special needs children.

Similarly, these measures are not intended for and are not endorsed for research outside of a Head Start evaluation context, even though such research may be an essential and appropriate prerequisite to program improvement. Many issues, involving technical and ethical considerations, best-use-of-resource concerns, protection of rights of subjects-and-family, have not been considered in this project for other than Head Start evaluations.

This project, however, can make an important contribution to research by providing data that will lead to new and potentially valuable hypotheses about the effects of programs -- hypotheses that will hopefully lead to basic research and the discovery of helpful knowledge through whatever measures and techniques the research community considers appropriate. This project can best contribute to such hypotheses through its measures being used to maintain a data base with systematic information on longitudinal program effects. Such a data base could yield important insights on relationships between characteristics of children at one age and at others.

Also, it is recognized that the evaluations contemplated in this project require information on independent variables, i.e., meaningful data on the various curricula, other services and other environmental impacts on the children in the program. And for a complete evaluation, measures of program effects on family, community and the institutions serving them would also be required. The measurement of these variables, however, is beyond the purposes of this project.

It should be pointed out, however, that the inclusion of data on these other variables in a data base, along with the data on children from these measures, could be particularly valuable. Such a combined data base would permit analysis of interrelationships and intra-relationships among children's entry status, family, community and institutional conditions, program variations, and program effects on family, community, institutions and children. It is, therefore, recommended that ACYF undertake measures-development efforts on these other variables and that their data be included in a compatible data base with the data to be derived from the measures of this project.

It should also be noted that ACYF has made significant progress in obtaining data on the independent variables through the Head Start Performance Indicators initiative. "Head Start Performance Indicators identify key dimensions of local program performance and service quality which can be measured objectively, are results oriented in nature, and include program services and activities" (Busa, 1980).

Objectives

The specific objective of this project is to develop a battery of measures of program effects on Head Start children between the ages of 3 and 7. These measures are to be used for the evaluation of Head Start programs while children are participating in Head Start and for longitudinal studies as children progress through first grade.

The principal objectives of this phase of the project were to (1) identify the outcomes to be measured, (2) develop a conceptual framework, (3) develop a set of recommended alternatives for ACYF on procedures for suitable measurement-development efforts, and (4) assist ACYF in the preparation of specifications for the selected measures. The ACYF specifications (DHEW, 1977) for the project define its objectives as follows:

> The objective of this project is to develop a battery of measures which will be used to evaluate the effectiveness of Head Start and Head Start related

programs in achieving greater social competence in children and to improve these programs. This battery shall encompass measures for each area of social competence as defined in the Head Start Performance Standards. This objective shall be achieved by a combination of development of new measures, adaptation of existing measures and the identification and inclusion of measures which are already suitable for use. When the battery is completed it shall be used in Head Start evaluations. It is, therefore, important to note that the development of the battery is to measure program effectiveness and not for individual assessment.

Since the intention of all future Head Start evaluations is to follow the children through at least their first year of school, the battery will be appropriate for use through at least the first grade.

The completed battery shall be developed on the basis of at least the following criteria:

1. reliability and validity for the target population;
2. satisfactory psychometric properties; especially the ability to discriminate between groups who have or have not had the treatment rather than discriminate between individuals;
3. ability to be administered by non-professionals;
4. ease of training for testers;
5. suitable for use with an age range from Head Start through first grade.

Basic Premises

Any undertaking in the development of measures contains both enormous opportunity and possible dangers. Used appropriately, the measures can provide valuable information for improving programs for children. Misused, the measures could lead to misleading and often dangerous assumptions about programs and children. It is, therefore, necessary to identify basic premises adopted in this project to guide future users on appropriate and inappropriate applications:

1. It is believed that any child can profit from a good developmental program experience, thus these measures are intended to assess the quality of the program experience through the assessment of children's developmental gains.

2. Because these measures are to be used in assessing the effectiveness of programs, the measures must be reliable for groups. That is, if the measure is administered a number of times over a short period, the "group mean" would remain stable. However, because a measure is reliable for a group does not mean that it is reliable for an individual. Therefore, these measures should not be used for the assessment of individual children.

3. These measures are derived from outcomes deemed of value or importance by Head Start parents, staff, and elementary school teachers and are intended to assess the effectiveness of the program in producing these outcomes. Therefore, the most legitimate comparison of the relative developmental gains between or among groups of children is in terms of the effectiveness of the program.

4. Because children differ in their rates and paths of development and previous experiences that affect their development, the relative effectiveness of programs should be compared only when they serve children with similar entry characteristics.

5. We believe in a developmental approach to programming. That is, it is the program's responsibility to adjust to the different developmental needs of children. Therefore, the measures should not be used as criteria for the placement of children in programs.

6. As a corollary to the above, it is appropriate to use these measures to make adjustments in the program.

7. We believe in the interrelatedness of children's cognitive, social-emotional and physical development. What affects one aspect of development affects the others. For example, if the health status of children is improved as a result of the health component of the program, one would also expect to find impacts in the other domains of development. Therefore, in order to measure the comprehensive impact of any program or component the measures from all domains should be used.

Uses of Project Outcomes

It is not the intent of this project to dictate the evaluation designs that should be used. The design selected depends upon the nature of the questions that are asked. However, the measures can be used to provide data relevant to questions addressing any of the following evaluation purposes:

1. To identify how well Head Start is achieving its national and local goals for children.

2. To determine the effectiveness of Head Start for children by comparing the development of those who have been in a Head Start program with those who are eligible but have not been in the program.

3. To determine the relative effectiveness of the various components of Head Start programs.

4. To identify successful programs so that their techniques may be disseminated to others.

5. To determine the effectiveness of various Head Start demonstrations and Head Start options.

6. To compare the effectiveness of Head Start with other child development programs that share common objectives.

7. To establish a meaningful basis for future evaluations that will permit comprehensive analysis of (a) the diversity in Head Start programs at a given point in time and (b) longitudinal analyses over time.

Non-Staff Personnel

The staff of the project has been assisted by a National Panel of 17 scholars and practitioners with expertise in many fields -- Head Start program operations, health and physical development, nutrition, special-needs children, cognitive and language development, social and emotional development, and psychometrics. Their racial/-ethnic backgrounds include Black, White, Hispanic, Asian and Native American. The names and affiliation of the panel are listed in Chapter II.

Also involved in this project are a large number of scholars and other child development specialists with expertise in the various areas of concern to this project. Their names also are noted in Chapter II. Many of their criticisms and recommendations proved to be invaluable in the development of this project.

An initial objective of this project was, and still is, the involvement of as many of those who are interested in Head Start as possible. This is to assure eliciting and consideration of the best possible alternatives, and also to develop a base of support and an audience for the outcomes reported by the measures finally developed.

Overview of Project Phases

This project has been planned to be carried out in three phases. Phase I comprises five major tasks:

1. Identification of the measurable characteristics indicative of children's development within the areas of Head Start objectives.

2. Development of a classification structure, encompassing all of the characteristics to be measured, to facilitate understanding and the usefulness of data on specific characteristics.

3. Identification of strategies, techniques and procedural plans through which an integrated system of measures can be developed for the required measures and structure.

4. Preparation of this technical report to document the findings and recommendations of the project and an options paper to present the Administration for Children, Youth and Families with alternatives for undertaking actual development of the measurement system.

5. Assistance to ACYF in the preparation of the specifications for measures, in accordance with the selected options, and in obtaining qualified contractors to carry out their development.

Phase II has four principal tasks:

1. Development and preliminary testing of techniques for measuring attainment of characteristics selected to provide comprehensive data as required to identify program effects within all areas of Head Start objectives.

2. Carrying out of a pre- and post-program pilot testing of component measures within each area, to evaluate and select those most suitable for the Head Start program effects measurement system.

3. Completion of documentation on selected measures.

4. Design of a national field test of the battery of selected measures for all areas, to be carried out during Phase III.

In Phase III, a national field test of the overall system to measure Head Start Profiles of Program Effects on Children will be carried out and evaluated.

Mediax Associates is the contractor for Phase I. During Phase II, Mediax will assist ACYF in the coordination and quality control of the measures-development efforts (tasks 1 - 3 above) being carried out by contractors selected for each of the major areas to be covered. Also, during Phase II, Mediax will prepare the design for the national field test. The Phase III effort will be carried out under a new contract.

The approaches used toward accomplishment of the initial three tasks of Phase I are described in general terms in Chapter II, and more detailed procedures and outcomes are set forth in succeeding chapters.

II.

METHODOLOGY

The methodology of the project includes general operating procedures followed throughout its development and other procedures addressed to specific tasks. Although particular procedures are referred to in other chapters, the complete outline is presented here.

General Operating Procedures

General operating procedures include the functioning of the project staff, relations with the National Panel and consultants, communications with the broad constituency of child-development specialists and organizations, dissemination, and relations with the national office of ACYF.

Staff

The central staff--consisting of the Director, Deputy Director and Technical Coordinator--conferred regularly at Mediax headquarters on all issues and procedures involved in the development of the project. The three non-resident staff members--Senior Scientist, Child Development Specialist, and Child Care Specialist--were involved in such conferences at many points, at the central office, at their respective offices, and by phone and correspondence. All major decisions of substance and procedure were initiated and directed by this staff personnel, whose identities are noted at the outset of this report.

National Panel

Early during the first phase of the project (March 1978), a National Panel of 16 child-development specialists and persons with related expertise (subsequently increased to 17) was formed to review theoretical statements and technical reports prepared by the staff, to prepare position papers on measurement in the several domains of child development, and to serve as a continuing resource on problems and issues confronted in the development of the project. They were selected by ACYF from 150 nominees solicited by the staff from several hundred organizations and individuals in the field of early childhood development.

Members of the National Panel include persons with expertise in child development, Head Start program operations and psychometrics; and they come from varying racial/ethnic backgrounds--Black, White, Hispanic, Asian and Native American.

The members of the National Panel and their institutional affiliations are listed below:

Saundra Aldridge, President, Head Start Staff Association

Eli Bower, University of California

Jean V. Carew, Stanford University

William Cross, Cornell University
Edward DeAvila, Linguametrics Group, Oakland, California
Rosalie Dixon, President, National Head Start Association
Dorothy H. Eichorn, University of California, Berkeley
Roger Givens, President, Head Start Parents Association
William Goodwin, University of Colorado
Susan G. Gordon, Presbyterian Hospital, Columbia University
Asa Hilliard, San Francisco State University
Luis Laosa, Educational Testing Service
Lee C. Lee, Cornell University
Ruth Marshall, President, Head Start Administrators Association
Winona Sample, California State Department of Health
Jean Collins Sanford, School of Public Health, Harvard University
Milton Silva, University of Wisconsin

This National Panel has thus far met in three, three-day sessions:

1. In Dundee, Illinois, June, 1978.
2. In Colorado Springs, Colorado, September, 1978.
3. In Burlingame, California, September, 1979.

In addition to these formal meetings of the Panel as a whole, informal conferences with sub-groups of the Panel were held from time to time at various places in the country, including Panelists with special expertise in the areas of Head Start operations, in children's development in health and nutrition, and in psychometrics. The central staff conferred on several occasions with individual Panelists, at their offices or at Mediax; and there have been frequent consultations with Panelists through correspondence and by phone. Most panelists also prepared comprehensive position papers on the assessment of child development in their respective areas of expertise.

Thus, the National Panel has been involved continually in the process of decision-making in the project. It will be involved also in the appraisal of proposals submitted by measurement-development contractors in response to RFP's, and in evaluating the procedures and products of those who are awarded contracts.

Consultants

Effort has been made throughout the project to solicit the ideas and judgments of a wide range of experts, in addition to those on the staff and National Panel. Some of them were involved through face-to-face interviews, others through correspondence and long-distance phone conversations, and a few of them were commissioned to prepare position papers on assigned problems and issues. Through these procedures, at least 70 consultants and National and regional ACYF staff have made substantive contributions to the project -- evaluating theoretical positions and procedures advanced by the staff, suggesting ideas and resources and techniques, and reviewing policy statements and reports. They include the following:

*Aguerre, Hank
Baer, Donald
*Baines, Hilton
Banks, James
Baumrind, Diana
Beller, Kuno E.
Bernal, Ernest
Bijou, Sidney
Bland, June
Blank, James A.
Block, Jack
Block, Jeanne
Bloom, Benjamin
Bronfenbrenner, Urie
Caldwell, Betty McDonald
*Carson, Linda
Cole, Michael
Comer, James P.
Cronbach, Lee J.
*Danavall, Elaine
*Digby, Fred
*Doyle, James P.
Faushel, Donald
*Fleischer, Roy
Forbes, Roy
Garcia, Eugene E.
*Garcia, John
Gella, Roland
Glass, Gene V
Gordon, Edmund W.
Guerrero, R. Diaz
Gussow, Joan Dye
Hartigan, John
*Hastings, Gerald
Howard, Rosanne B.
Huerta, Alfred
Jacobson, Elaine
*Jordan, John
Juhas, Loretta
Kagan, Jerome
Katz, Lilian
Kearsley, Richard
Lipsitt, Lewis P.
Lord, Frederick
*Miller, Samuel
Moore, Evelyn
Morissey, Patricia Garland
Murphy, Lois
Parges, Stephen
Patterson, Claude
Pfeiffer, Anita B.
*Phillips, Margaret
Ramirez, Judith
*Randolph, Linda
Ryle, Anne L.
Sawin, Douglas B.
Shipman, Virginia
Skinner, B.F.
Slatt, Carl
Sroufe, Alan
Swaminathan, Hariharan
Thomas, John
*Vincinanza, Paul
Walker, Deborah
Wenar, Charles
White, Burton L.
White, Sheldon
Wright, Benjamin
Yarrow, Leon J.
Zigler, Edward
Zimiles, Herbert

* ACYF National and Regional Staff

It should be noted that listing these consultants does not imply their endorsement of the theoretical positions and procedures adopted by the staff and National Panel. Indeed, a few of the consultants contributed importantly to the project by sharply criticizing -- and thus forcing more thorough analysis of -- positions eventually taken.

It should be noted also that the above list does not include 375 Head Start parents and staff members and K-2 teachers, together with 27 facilitators, who participated in the series of Input Workshops conducted in the several ACYF regions. They, too, served as very important consultants to this project.

Dissemination

Since this project has important implications not only for Head Start, but also for the wider range of theoreticians, researchers and practitioners in the field of early childhood development, effort has been made from the outset to disseminate all major developments to a broad "constituency". Several approaches to this end have been used.

An "Overview of the Project" was mailed early to several hundred organizations and individuals in the early child-development field, including those who were invited to nominate persons for the National Panel.

Members of the staff have interpreted the project through presentations at the following conferences of professional organizations:

National Head Start Annual Meetings:

Cleveland, Ohio, January, 1978
Denver, Colorado, May, 1979
Miami, Florida, May, 1980

National Head Start Board Meeting, Kansas City, Missouri, February, 1979.

Northeastern Head Start Directors Association Meeting, Boston, Massachusetts, April, 1979.

National Association for the Education of Young Children, Annual Conference, New York, N.Y., April 1978.

American Educational Research Association, Annual Meeting, San Francisco, California, April, 1979.

Interviews and correspondence with the consultants listed above have served as an important means of dissemination. So, also has correspondence with scores of individuals and organizations who have written to the staff for information about the project.

Three issues of the project's NEWSLETTER have been published and distributed--in June 1978, in November 1978, and in January 1980. Each issue was mailed to approximately 800 individuals and organizations, carrying information about the project -- objectives, planned stages of development, theoretical problems and issues involved, major procedural steps, preliminary reports on the findings of investigations, personnel changes, and related matters.

The general purpose of these efforts at dissemination is to acquaint individuals and organizations in the child-development field throughout the country with the objectives, nature, implications and progress of the project, and also to invite their suggestions and criticisms.

Relations with ACYF

Every major theoretical position and procedural step taken by the project has been discussed in advance and approved by the ACYF Project Officer; and formal reports on progress have been submitted to her regularly. The staff is in continual communication with this official, through conferences in Washington or at Mediax, and has found her to be extremely helpful in guiding the project.

On at least three occasions during the project, the staff has discussed accomplishments and perspectives in meetings with key national officials of ACYF, including the Commissioner and Project Officer. Discussions at these meetings served to determine the direction the project would take at several points where critical decisions were necessary.

Summary

Thus, the general operating procedures of this measurement-development project have involved the participation of a very large number of persons -- notably the central staff, the 17-member National Panel, nearly 500 consultants (including Head Start and related personnel in the field), several hundred child-development specialists and organizations throughout the country, and key personnel of the ACYF National and Regional Offices. Their input has effectively influenced the course of the project's development.

Defining the Issues

An early undertaking of the project was to review and analyze relevant literature for the purpose of determining the state-of-the-art -- contesting theoretical positions on the nature and measurement of early childhood development, available measuring instruments, and technical problems and issues. Reports of the main two precursors of the project --Anderson and Messick, Social Competence in Young Children (1973), and Raizen and Bobrow, Design for a National Evaluation of Social Competence in Head Start Children (1974) --were studied, as was Head Start's official program guide, Head Start Performance Standards (1975). IN addition to these an extensive library of relevant materials was assembled --published books, journal articles, reports of Head Start evaluation studies, measuring instruments, and unpublished materials obtained mainly from consultants and other correspondents.

These materials were analyzed for the purpose of identifying points of view, approaches and data relevant to the two major tasks of the project: (1) to define the developmental charcteristics of young children to be measured as a basis for appraising program effects, and (2) to develop recommendations for measuring the development of these characteristics in children 3 to 7 years old.

The main procedural thrusts of the project were determined in light of this analysis of relevant literature. In order to facilitate discussion and decision-making within the staff and among members of the National Panel, the main findings of this analysis were compiled and/or interpreted in a series of in-house documents, chief among which are the following.

1. "Panelist Guide", January 1978 (Mediax, 1978h).

2. "Measures of Development Among Young Children: Socio-Emotional Domain," by Lilian Katz and Elaine Jacobson, May 1978.

3. "Resource Reference Book: Test Collection, Educational Testing Service," Vol. 1, June 1978 (Mediax, 1978l).

4. "Selected Current Literature Series: Dimensions of Child Development," Vol. 2, Nos. 1 to 32, June 1978 (Mediax, 1978m).

5. "ERIC Literature Indexed to Dimensions of Child Development," Vol. 3, June 1978 (Mediax, 1978e).

6. "Compendium of Information Gathering Techniques Utilized in Previous Head Start Research Projects," Vols. 1 and 2, June 1978 (Mediax, 1978a).

7. "Current Issues in the Development Project for Head Start Profiles of Program Effects on Children", August 1, 1978 (Mediax, 1978c).

8. "Measuring the Effect of Head Start Programs on Future Academic Achievement" (Ryle, 1978)

9. "Assessing Head Start's Impact on Cognitive Development of Children" (Zimiles, 1978)

Also of major importance in this regard are the "positon papers" prepared by members of the National Panel with reference to their respective areas of expertise. These reports are listed below.

1. "Assessment of Personal and Social Growth Factors Through Play" (Bower 1978).

2 "Head Start Profiles of Program Effects on Children" (Carew, 1978).

3. "Recommendations of Dimensions to be Measured in the Area of Nutrutional Development of 3 to 7 Year Olds"(Collins-Sanford, 1978).

4. "Assessment of Emotional Development in Children Participating in Head Start Programs" (Cross, 1978).

5. "Review of Tests Available for Piagetian Conservation Tasks" (DeAvila, 1978).

6. "Head Start Profiles of Program Effects on Children: Motor Development" Eichorn, 1978).

7. "The Measurement of Children's Motivation as a Program Outcome" (Goodwin, 1978).

8. "Head Start Profiles of Program Effects on Children -- Medical and Dental Domain" (Gordon, 1978).

9. "Evaluating the Effects of Head Start on Children's Cognitive Development" (Laosa, 1978).

10. "Assessing Communicative Competence via the Repertoire Model" (Lee, 1978).

11. "I'm back!!! Children's Social Behavior: Assessment Considerations and groups in the seven workshops. A trained "facilitator" of the same racial/ethnic background worked with each group.

12. "Teaching Head Start Children Surival Skills in the Mainstream Society" (Silva, 1978).

Determining Child Characteristics To Be Measured

Three main approaches were followed in determining the characteristics of children to be mesured -- a survey of child-development scholars, a series of workshops with Head Start and related personnel in the field, and an analysis of characteristics identified in relevant literature.

Scholar's Survey

A survey was made among selected early-childhood scholars to determine what developmental characteristics they consider important for evaluating the effects of Head Start and similar programs. Dr. J. McVicker Hunt, senior scientist on the project, wrote to 31 scholars whom he considered representative of varying points of view, asking, among other things, that they identify the "domains or broad relms of behavior you have found to be most significant and useful for evaluating the quality, positive or negative, of the development of children aged from 3 to 7 years". 3. He also asked that they specify "those behavioral characteristics which define operationally the quality of development within these domains and which you feel might be useful in assessing or measuring them".

The 22 scholars who responded to the inquiry suggested 271 more or less discrete child characteristics for use in measuring program effects on young children. They were classified by the staf in 18 dimensions of development in 4 broad categories: Cross-Domain Affective Attainments, Physical and Health Development, Cognitive Development, and Social Development.

The procedures and findings of this survey of scholars were compiled and analyzed in an in-house document entitled "Preliminary Interim Report on Child-Development Attainments Identified by Scholars for Evaluating Program Effects," April 10, 1978, (Mediax, 1978i).

Head Start Input Workshops

A field study was made to determine what attainments (or "characteristics") Head Start parents and staff members, together with K-2 teachers, think Head Start should help children to develop. For this purpose, a series of seven, two-day workshops were conducted during April and May 1978 at six locations (two at one site) in different parts of the country. Participants were selected by Head Start grantee agencies in cooperation with all 11 of the ACYF regional offices. Effort was made to include representatives of different racial/ethnic groups among particpiating parents, Head Start staff members, and K-2 teachers.

The seven workshops were conducted in the ACYF regions and locations listed below:

Regions I-II-III: New York, N.Y., April 12-13, 1978
Region IV: Atlanta, GA., April 19-20, 1978
Regions VI-VIII: Dallas, TX., April 25-26, 1978
Region VI: Dallas, TX., April 27-28, 1978
Regions VIII-IX-X: San Francisco, CA., May 2-3, 1978
Region VII: Kansas City, MO., May 11-12, 1978
Region V: Chicago, IL., May 17-18, 1978

There were 375 workshop participants, They were distributed by role as follows:

Parents	187
Head Start Staff	139
Public School K-2 Teachers	62

Some represented the dual role of parent and staff, which explains why the above frequencies exceed the total of 375.

The 375 participants were distributed by racial/ethnic background as follows:

Black	35.7%
White	36.0%
Hispanic	18.9%
Asian or Polynesian	4.8%
Native American	4.5%

Table II provides a further breakdown of participants by region and living environment.

The procedures followed in organizing and conducting these workshops are reported fully in the staff document entitled "Report on Process of Input Workshops", May 27 (revised August 15), 1978 (Mediax, 1978k). Briefly, they include the following:

1. In preparation for the workshops, a pilot test of draft materials was conducted with 48 Head Start parents, staff members and K-2 teachers in a community near the mediax offices. They were presented the 18 categories of attainments identified by scholars, and asked to list specific competencies that Head Start should help children develop for each of them. They were also asked to add other desirable areas of development not covered by the specified constructs.

   Analysis of responses to this pilot test led to the refinement of several of the original 18 items and to the addition of 8 more, making a total of 26 discussion questions.

TABLE II

NUMBER OF PARTICIPANTS IN INPUT WORKSHOPS BY DEMOGRAPHIC GROUP

WORKSHOP	ALL	BLACK	HISP	NATIVE	ASIAN	WHITE	PARENT	TEACHER	STAFF	MIGRANT	RURAL	URBAN	NE	SE	CENT	SW	W
All Workshops																	
*Day 2	364	131	67	17	16	133	183	62	132	16	126	222	82	50	93	103	36
Others	11	3	4	0	2	2	4	0	7	0	4	7	0	0	6	3	2
Total	375	134	71	17	18	135	187	62	139	16	130	229	82	50	99	106	38
Atlanta																	
*Day 2	50	26	0	0	1	23	25	7	19	0	15	35	0	50	0	0	0
Others	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Total	50	26	0	0	1	23	25	7	19	0	15	35	0	50	0	0	0
Chicago																	
*Day 2	53	18	10	0	0	25	24	16	13	0	17	36	0	0	53	0	0
Others	3	1	1	0	0	1	0	0	3	0	2	1	0	0	3	0	0
Total	56	19	11	0	0	26	24	16	16	0	19	37	0	0	56	0	0
Dallas One																	
*Day 2	60	21	17	11	0	11	29	10	25	2	44	14	0	0	2	58	0
Others	1	0	1	0	0	0	1	0	0	0	1	0	0	0	0	1	0
Total	61	21	18	11	0	11	30	10	25	2	45	14	0	0	2	59	0
Dallas Two																	
*Day 2	45	12	20	3	0	10	25	1	22	14	10	21	0	0	0	45	0
Others	2	0	2	0	0	0	2	0	0	0	0	2	0	0	0	2	0
Total	47	12	22	3	0	10	27	1	22	14	10	23	0	0	0	47	0
Kansas City																	
*Day 2	14	10	0	1	0	3	8	2	5	0	3	11	0	0	14	0	0
Others	3	2	0	0	0	1	1	0	2	0	1	2	0	0	3	0	0
Total	17	12	0	1	0	4	9	2	7	0	4	13	0	0	17	0	0
New York																	
*Day 2	82	37	10	0	7	28	40	16	30	0	15	67	82	0	0	0	0
Others	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Total	82	37	10	0	7	28	40	16	30	0	15	67	82	0	0	0	0
San Francisco																	
*Day 2	60	7	10	2	8	33	32	10	18	0	22	38	0	0	24	0	36
Others	2	0	0	0	2	0	0	0	2	0	0	2	0	0	0	0	2
Total	62	7	10	2	10	33	32	10	20	0	22	40	0	0	24	0	38
*Demographic Group Overlap																	
All	364	131	67	17	16	133	183	62	132	16	126	222	82	50	93	103	36
Black	131	131	0	0	0	0	72	18	47	0	31	100	37	26	31	33	4
Hispanic	67	0	67	0	0	0	34	11	24	15	20	32	10	0	17	37	3
Native American	17	0	0	17	0	0	10	1	8	0	14	3	0	0	3	12	2
Asian	16	0	0	0	16	0	7	1	8	0	0	16	7	1	0	0	8
White	133	0	0	0	0	133	60	31	45	1	61	71	28	23	42	21	19
Parent	183	72	34	10	7	60	183	1	12	8	62	113	40	25	44	54	20
Teacher	62	18	11	1	1	31	1	62	0	2	19	41	16	7	25	10	4
Staff	132	47	24	8	8	45	12	0	132	7	50	75	30	19	25	46	12
Migrant	16	0	15	0	0	1	8	2	7	16	0	0	0	0	0	16	0
Rural	126	31	20	14	0	61	62	19	50	0	126	0	15	15	34	52	10
Urban	222	100	32	3	16	71	113	41	75	0	0	222	67	35	59	35	26
NE	82	37	10	0	7	28	40	16	30	0	15	67	82	0	0	0	0
SE	50	26	0	0	1	23	25	7	19	0	15	35	0	50	0	0	0
Cent	93	31	17	3	0	42	44	25	25	0	34	59	0	0	93	0	0
SW	103	33	37	12	0	21	54	10	46	16	52	35	0	0	0	103	0
W	36	4	3	2	8	19	20	4	12	0	10	26	0	0	0	0	36

*Includes only participants who completed the questionnare on the second day of the workshop.

Twenty-four of the discussion topics has a common beginning phase followed by a descriptor relevant to one particular area of development, as follows:

Describe those things you want a child to be able to do that prove that he or she:

1. is a good listener.
2. speaks well.
3. can solve problems and reason well
4. is good with numbers and has arithmetic skills.
5. has reading readiness skills.
6. has good dental habits.
7. has good health and safety habits.
8. has good physical development and is able to use whole body skillfully.
9. is able to use hands and fingers skillfully.
10. shares with other children.
11. has good eating habits/knows about food.
12. is able to use senses (hearing, seeing, feeling, smelling).
13. takes on complete tasks.
14. expresses feelings and emotions.
15. is sensitive to feelings of others.
16. gets along with other children.
17. gets along with adults.
18. is able to play well, including acting-out adult and fantasy roles.
19. knows rules and customs of society.
20. has school readiness skills.
21. reads well.
22. understands things about nature.
23. understands about the things in the world that humanity has created.
24. feels good about himself or herself.

The other two discussion topics were:

25. We are living in a society of many nationalities and races:

    a. Is there anything you want your children to learn or be able to do to help them get along with other races or nationalities?

    b. Is there knowledge or skills particular to their own race or ethnic gorup that you want them to have?

26. List any other skills, knowledge, attitudes or habits not covered in the above that you would like children to have. In each case, describe what children would be able to do that proves attainment.

2. At each of the seven regional workshops previously listed, participants were organized into role groups (i.e., parents, staff, K-2 teacher) of from four to seven persons each, the members of each group (with two or three exceptions) being of the same racial/ethnic background. There was a total of 78 of these groups in the seven workshops. A trained "facilitator" of the same racial/ethnic backgorund worked with each group.

3. On the first day of the workshop, each of these small role groups was presented with the 26 open-ended "Discussion Topics". The topics were presented so as to elicit participants' judgments concerning the specific characteristics desirable

for Head Start and school to help children aged 3 to 7 to develop in a particular area. Each topic was accompanied by a relevant "Wrap-Up Issue" designed to elicit further discussion on the area.

4. Each small workshop group selected a chairperson, and decided the sequence in which its members wanted to react to the discussion topics. Before proceeding with the discussion on any topic, the chairperson polled the group to determine how many members considered the area to be discussed as "valuable" for their children. Responses for each topic were recorded with a magic marker on a large flipsheet mounted on an easel. Each different response from group members was recorded, even if it represented the judgment of only one participant.

5. The desirable characteristics listed by the several workshop groups on the first day were compiled overnight, and presented the next morning in the form of a "Consensus Questionnaire" to the workshop as a whole. Each individual participant was asked to rate each characteristic as "Harmful", "No Value", "Minimum Value", "Some Value", "Important", or "Absolutely Essential".

Participants in these Input Workshops identified approximately 13,000 statements. These were used to generate 4,896 workshop items that were rated on Day 2. The final listing resulted in 1,516 more or less discrete characteristics that Head Start should help children develop. They were compiled in two staff documents entitled "Report of Attainments Identified in Input Workshops" and "Composite List of Developmental Objectives (Coded) for Children Aged 3 to 7", both of which were prepared in December, 1978.

It should be noted that the participants were not selected as a stratified random sample of Head Start parents, staff, and related K-2 teachers; and, hence, their judgments do not provide a basis for firm generalizations about the populations from which they were drawn. It is probable, however, that the range of characteristics they consider desirable for children aged 3 to 7 closely approximates the range one would find from a statistically representative sample.

Literature Search

The assembled body of relevant literature previously noted was analyzed for the specific purpose of discovering desired developmental characteristics of young children different from those identified in the Scholar's Survey and the Input Workshops. Only 191 different characteristics were noted (about 1.0 per cent of the total number identified). All of these additional characteristics were classifiable in areas already covered.

Analysis of Findings

The findings of the Scholar's Survey, Input Workshops and literature search were discussed extensively by the staff and members of the National Panel, and were analyzed in great detail. Resulting therefrom is the classified list of children's developmental characteristics presented and interpreted subsequently in this report. They are the desirable outcomes of the Head Start program for which this project undertakes to select or develop measures.

Analysis of Available Measures

A recent review of developmental assessment instruments estimated that there may be some 5,000 measures available for assessing development in various characteristics of young children (International Training Consultants, 1978). This review, which focused on the use of measures by Head Start teachers and staff, concluded that approximately 180 of them are so used. Since a large proportion of these measures are not suitable for program evaluation purposes, an effort was made as part of this project to identify those used in Head Start research and evaluation. Seven studies were looked at in some depth. They are listed below, with the particular report or reports cited that were reviewed.

The projects examined are the following:

1. ETS-Head Start Longitudinal Study -- Educational Testing Service: Shipman, Boroson, Bridgeman, Gant & Mikovsky (1976).

2. National Day Care Center Study -- Abt Associates: Abt Associates (1977); Ruopp, Travers, Glantz, & Coelen (1979).

3. National Survey of Head Start Graduates and Their Peers -- Abt Associates: Royster & Larson (1978).

4. Study of the Head Start Handicapped Effort -- Applied Management Sciences: Rader, Thouvenelle, Vogel & Davis (1978); Vogel & Rader (1978).

5. National Home Start Evaluation -- High/Scope Educational Research Foundation and Abt Associates: Love, Nauta, Coelen, Hewett & Ruopp (1976).

6. Head Start Planned Variation Evaluation -- Huron Institute: Walker, Bane & Bryk (1973).

7. Project Developmental Continuity Evaluation -- High/Scope Educational Research Foundation: Love, Wacker & Meece (1975); Love, Granville & Smith (1978).

In addition to examining these reports directly, Mediax staff drew upon the reviews of others. These included:

1. System Development Corporation (1972): review of measures used in the 1967-68 evaluation.

2. Raizen & Bobrow (1974): review of measures included in the Rand Corporation's "design for a national evaluation of social competence in Head Start children".

3. Office of Child Development (1975a): listing of final reports of Head Start evaluation projects.

4. Datta, McHale & Mitchell (1976): summary of 1966-1969 Head Start evaluations.

5. Mann, Harrell & Hurt (1977): review of Head Start research since 1969.

6. Educational Testing Service's Head Start Test Collection: compiled in a document entitled "Resource Reference Book-Test Collection, ETS, Vol. I, June 1978 (Mediax, 1978l).

Descriptions of the measures used in these studies, together with relevant findings, were compiled in the project document entitled "Compendium of Information-Gathering Techniques Utilized in Previous Head Start Research Projects" (July 5, 1978). The final list included a varied assortment of measuring techniques--paper-and-pencil tests, inventories, rating scales, medical records, observation scales, interaction procedures, and sociometric devices.

Several comprehensive reviews of measuring instruments were also examined for the purpose of identifying additional techniques, whether or not they had been used in Head Start evaluations. These included reviews by Buros (1965, 1972); Butler, Gotts, Quissenberry & Thompson (1971); Hoepfner, Stern & Nummidal (1971); Johnson (1976); Johnson & Bommarito (1971); Robinson & Shaver (1974); and Walker (1973).

A special analysis was made by project consultants Lilian Katz and Elaine Jacobson of measures used to assess social and emotional development: "Measures of Development Among Young Children: Socio-Emotional Domain", (Katz & Jacobson, 1978). A number of papers prepared by members of the National Panel also identified and reviewed instruments in the various domains. These included measures for the cognitive domain (Laosa, 1978), measures of children's motivation (Goodwin, 1978), nutrition measures (Collins-Sanford, 1978), measures of personal and social growth (Bower, 1978), measures of emotional development (Cross, 1978), and Piagetian measures (DeAvila, 1978). Panel members and other experts were also asked to suggest relevant measures.

Through these procedures a total of more than 200 measures were identified. The next task was to appraise them systematically for potential use. This appraisal was based on several criteria--the nature of the program objectives for which the measure was recommended and used, the ages of the children to whom the measure had been administered, and the sensitivity of the measure to program effects. A detailed discussion of the criteria used and the results of the analysis are presented in Chapter V.[1] These analyses yielded conclusions consistent with those of other reviewers and critics; namely, that very few measures show content validity as defined by the developmental characteristics of children identified through the Input Workshops. Information on construct validity is virtually nonexistent. Moreover, very few measures possess sufficient reliability to warrant confidence in evaluation findings based on their use.

---

[1] Since this review and analysis was completed in 1978, we have attempted to keep abreast of potential developments in Head Start evaluations. Thus, the conclusions reported in Chapter V include information from more recent studies, including reports of the Home Start Follow-Up Study, Parent Child-Development Center evaluation, Child and Family Resource Program evaluation, the evaluation of the ACYF bilingual curriculum development projects, and the Basic Educational Skills program evaluation.

Although the examination of these measures revealed that some of them were very well conceived and characterized by acceptable psychometric properties, even they failed to comprehend the range of child-development characteristics shown by this project to be valued by scholars and by parents and teachers in the Head Start programs to be evaluated. Thus, it became clear to the staff and National Panel that new measures would have to be developed to adequately service the purposes of this project. Needed is a comprehensive and coherent battery of measures geared specifically to the several domains and dimensions of child development defined as the basis for appraising program effects on children.

III

CONSIDERATIONS IN DEFINING CHARACTERISTICS

The selection of characteristics to be measured in a program's evaluation is most critical to the evaluation and, also, to the program's future content. Following is a discussion of the most important considerations which have guided the selection process in this project:

- Need for a comprehensive definition of Head Start program objectives.
- Need for valid indicators
- Diversity
- Problem of defining domains
- Problems in classifying child development characteristics
- Criterion, background and other variables

Need for a Comprehensive Definition of Program Objectives

The task of selecting developmental characteristics of children that are to be measured in this project is essentially that of identifying the goals of Head Start. The starting point of measurement and evaluation is what the program seeks to accomplish, its objectives. The task is not to define what the program's objectives should be, which is a function of others; rather, given statements of goals, it is to identify those specific measurable characteristics that evidence their attainment.

The Head Start Performance Standards (1975b) defines the program's goals, as:

> ...to bring about a greater degree of social competence in children of low income families. By social competence is meant the child's everyday effectiveness in dealing with both present environment and later responsibilities in school and life. Social competence takes into account the interrelatedness of cognitive and intellectual development, physical and mental health, nutritional needs, and other factors that enable a developmental approach to helping children achieve social competence (p.1).

A basic task of this project is, therefore, to define the measurable characteristics that evidence attainment of the goal of social competence. As noted in the Introduction to this report (p. 15), the Educational Testing Service undertook this task in 1972. The 29 "facets of social competence" they identified are listed in Table III

The somewhat similar project undertaken by the Rand Corporation in the following year identified 42 outcome variables associated with "social competence" as listed in Table IV.

TABLE III

Facets of Social Competency
(Anderson and Messick, 1973)

1. Differentiated self-concept and consolidation of identity
2. Conception of self as an initiating and controlling agent
3. Habits of personal maintenance and care
4. Realistic appraisal of self, accompanied by feelings of personal worth
5. Differentiation of feelings and appreciation of their manifestations and implications
6. Sensitivity and understanding in social relationships
7. Positive and affectionate personal relationships
8. Role perception and appreciation
9. Appropriate regulation of antisocial behavior
10. Morality and prosocial tendencies
11. Curiosity and exploratory behavior
12. Control of attention
13. Perceptual skills
14. Fine motor dexterity
15. Gross motor skills
16. Perceptual motor skills
17. Language skills
18. Categorizing skills
19. Memory skills
20. Critical thinking skills
21. Creative thinking skills
22. Problem-solving skills
23. Flexibility in the application of information-processing strategies
24. Quantitative and relational concepts, understandings and skills
25. General knowledge
26. Competence motivation
27. Facility in the use of resources for learning and problem-solving
28. Some positive attitudes toward learning and school experiences
29. Enjoyment of humor, play and fantasy

TABLE IV

Outcomes Associated with Social Competence

(Raizen and Bobrow, 1974; numbers and rephrasing by Mediax)

Health and Nutrition

1. Reduced incidence of disease
2. Reductions in iron deficiencies
3. Reduction in protein deficiencies
4. Improved dental health
5. Remediation of visual impairment
6. Remediation of hearing impairment
7. Good overall health (through any in-depth health assessment)
8. Gains in growth and stature
9. Increased vigor

Perceptual-Motor, Cognitive and Language Skills

10. Visually guided fine motor
11. Visual recognition and discrimination
12. Cognitive functioning
13. Problem solving
14. Numerical readiness
15. Letter-number recognition
16. Vocabulary (English & Spanish)
17. Comprehension (English & Spanish)
18. Oral language
19. Inter-personal communication
20. Test-taking behavior

Social & Personal Development (recommended for first year of public school only)

21. Role behaviors toward peers and their responses
22. Role behaviors towards teachers and their responses
23. Role behaviors toward parents and their responses
24. Executive skills
25. School success
26. Adjustment to school
27. Direction following and task completion
28. Goal-setting and self-evaluation
29. Intentional-incidental learning
30. Reinforcement style
31. Curiosity
32. Perception of spatial perspective
33. Perception of situational perspective
34. Perception of cultural perspective
35. Boundary elasticity
36. Multiple solutions
37. Barrier behavior
38. Consequential reasoning
39. Conflict-resolving alternatives
40. Attitudes toward school
41. Attitudes toward intellectual achievement
42. Attitudes toward self

Although the outcomes identified through the ETS and Rand projects include a few fairly specific characteristics (e.g., reduction in iron deficiency), most are expressed in general terms that are not directly measurable. (e.g., differentiated self-concept and consolidation of identity, morality and prosocial tendencies, adjustment to school, general knowledge, etc.). They leave undefined the specific behaviors that must be measured to evidence the attainment of desired characteristics. They are useful, however, in further defining the broad areas of desired characteristics, the subconstructs of the global construct of desired outcomes that is social competence.

This project adopted the view that the use of scholars to initially identify broad areas of desired Head Start outcomes, as was carried out by ETS and Rand, is both necessary and desirable; but that, in consideration of bias and diversity issues discussed below, the specific characteristics to be measured must be identified by Head Start parents, staff and public school teachers.

Need for Unbiased Indicators

Just as ETS and Rand resorted to the use of descriptors of areas of desired characteristics, i.e. subconstructs, to communicate desired Head Start outcomes, so will this project and the battery of measures to be developed. Ideally, then, we should undertake to measure such subconstructs directly; i.e. to obtain direct data on children's attainment of the overall qualities represented by the particular subconstructs. But the judgments that are necessarily involved in obtaining data on such "overall qualities"--sensitivity, language skills, memory skills, general knowledge, self-concept, good health--are highly subjective. They are influenced by the particular rater's bias on what are the most important elements in the desired overall quality. And, because a direct measure of many of these overall qualities would require a judgment by someone intimately familiar with each child being measured, such measurement is subject to influence by many personal factors incompatible with the requirement of this project for objective measures. (Such overall judgments are also highly unreliable in that they require the evaluation and integration of a large number of individual behaviors which may vary greatly from observation period to observation period.)

What is needed is the identification of more objectively measurable indicators of the attainment of the qualities represented by these subconstructs, indicators that can be recognized and interpreted by trained observers. The particular method of aggregating data on a set of indicators for each subconstruct must, however, be separately validated. Such validation will necessarily also involve data representing judgments on the subconstruct level.

Each of the indicators of the attainment of the desired qualities represented by a subconstruct should be somewhat representative of what is overall desired. As a totality, all the indicators (i.e., all the characteristics of children selected for measurement in the battery) should encompass all the qualities represented by the various subconstructs and by "social competence". Indicators representing cultural and regional differences of desired outcome qualities must, therefore, also be included; otherwise, the set of characteristics selected for measurement will be inherently biased.

But measurement of the attainment of desired outcome qualities is not sufficient to determine program effects. Progress of children toward such attainments is also important. Therefore, characteristics have to be identified that are indicative of the status of children at the beginning of Head Start programs and at other points, as they progress in their development in each of the areas of concern. These must also encompass cultural and regional differences and, therefore, require input from persons familiar with the meaning of behaviors of the particular children being measured.

Thus, a sufficient definition of characteristics to be measured must cover three requirements for data, (1) on the various subconstructs defining social competence, (2) on the set of desired outcome characteristics in the area of each subconstruct, and (3) on the characteristics representing entry and intermediate statuses of children as they develop toward the desired outcomes. Each must involve the judgments of persons familiar with the patterns of behavior of the particular children being measured to avoid bias. The selection of the characteristics necessary to be included in the final battery of program effects measures, out of all those initially identified as relevant, is a task for the measures-development contractors during Phase II of the project.

Diversity

Most measures for young children are designed around a set of expectations of what children should be capable of doing at a particular age. Specific capabilities are usually selected as norms for the age based on data obtained on a "representative" population. There are few instances in early child development, however, where groups of children should be expected to acquire particular behaviors at a specific age. Such single standards may be appropriate for all children completing a formal training stage, e.g. upon graduating from high school, but they are most inappropriate for young children in early or middle stages of development; they are also most inappropriate as criteria of Head Start program effects.

Children with different background experiences, of necessity, pursue different paths of development. At any point in time, such as when entering Head Start, they will, therefore, have attained different characteristics in the various areas of their development. Consequently, the paths of their development, as nurtured by Head Start, the public schools, and other influences, should also be different to elicit the series of developing characteristics that will ultimately lead to desired outcomes. And there will be further diversity in the outcomes desired by various communities, parents and, increasingly with their years, the children themselves.

A similar view has been expressed in a review of the Rand report by "Black and Spanish surnamed professionals":

> Since each community or each minority may have different values and therefore different ideas about what constitutes competent behavior, investigating a set of social competence variables across sites is inappropriate...most of the tests considered have been standardized on and measure skills of middle-class urban children...Present measures do not take into account minority culture styles and values (Raizen & Bobrow, 1974, pp. 422, 423)

The Head Start Performance Standards recognize the diversity in values, priorities, and needs within the overall Head Start population, by prescribing local determination of objectives and curriculum. The goal of Head Start, social competence, is defined in terms that allow for diversity -- "the child's everyday effectiveness in dealing with both present environment and later responsibilities in school and life." It is necessary that these measures also accommodate the differences in children's "present environments", "later responsibilities", "school" programs, and chosen "life".

The National Panel of this project recommended that measures developers be required to separately validate all measures on Black, White, Hispanic, and one other ethnic/racial group of children, to assure the accommodation of diversity and the elimination of bias. This recommendation does not imply that there are only four "different" groups among Head Start, or that any one of the specified groups does not have important differences among subgroups. The recommendation is based on the presumption that any test battery that accommodates the diversity in the specified four groups will also accommodate the range of differences that exist in the overall Head Start population.

The battery of measures to be developed must allow in its design for such diversity in the paths and indicators of children's development. The specific characteristics to be identified for possible measurement, therefore, must also encompass these differences; they must be obtained from those persons most familiar with the different mainfestations of desired development in the particular children being measured.

Problem of Defining Domains

In most achievement test construction, whether norm referenced or objective referenced, the criterion for item selection is based on a conception of a set of characteristics representing a desired status at the end of the program. In the traditional norm-referenced approach, items are selected that discriminate between individuals around these indicator characteristics. In the objective referenced approach, also, each item is selected to represent end-of-program status. Both of these approaches, therefore, result in measures that are not sensitive to characteristics of children at the beginning of a program, especially where such children enter programs with backgrounds out of the mainstream for which tests are developed.

Many critics of traditional testing have pointed out that if we are to truly measure achievement, or program effects on achievement, we must first have a well-defined domain over which the child is achieving. The concept of achievement, and program effects in this project, implies locating the status of children in a domain at the beginning of a program and locating these same children in the domain at another point in time, such as the end of the program. The change in status over the time of the program is really what program effects are all about.

In contrast, measuring children in terms of their attainment of a set of characteristics representing only a desired status at the end of a program, as is inherent in using most available measures, tells little about change in status over the domain. Linn (1979, p. 2) argues that the use of such tests "results in more of an aptitude measure than an achievement measure". To effectively measure program effects, we must identify the characteristics defining each domain over which we desire to measure change.

There are problems in comprehensively defining domains. Linn (1979, p.5) points out that:

> ... indeed, some would argue that it (i.e. defining domains) is a hopeless task for other than trivial content items ... but the difficulty of developing tests that accomplish the goal of measurement is not an adequate reason for abandoning the goal in favor of a different one that can be more readily achieved.

The approach adopted in this project to defining domains involved analysis of a universe of desired characteristics hypothesized to encompass those defining the various domains. The universe was compiled through statements of indicators of desired development in each of the areas of concern to Head Start, elicited from Head Start parents, staff and public school teachers. The rationale for this approach is that such persons could best articulate the appropriate and diverse indicators of desired develoment in the various areas targeted by Head Start. And if a sufficiently large sample of characteristics is elicited from an adequate number of representatives from each of the diverse populations, the complete range of desired developmental characteristics will be generated from which the various domains and other constructs may be defined. The procedures used in implementing this approach are described in the methodology chapter of this report.

It should be noted that the approach here adopted does not lead to a universally applicable definition of domains. It is only claimed to lead to definitions useful to Head Start. The measures based on these procedures, therefore, may be assumed to be valid for measurement of the effects of Head Start or other programs with similar populations. Measures for programs with different populations would require an identification and redefinition of domains based on procedures that are appropriate for their own populations.

Also, although the procedures followed are believed to adequately identify the variety of characteristics defining all of the domains of interest, the relationships of these characteristics to the domains and to each other requires confirmation. Data obtained through the pilot and field-testing of the measures in all the domains, and data from subsequent applications of the selected battery, should be analyzed to test and, when necessary, reformulate these hypotheses. Such analysis should also result in an identification of redundancies in the measures and, consequently, a reduction in the total number of characteristics required in the battery.

Problems in Classifying Child Development Characteristics

The process of analysis of a universe of desired child development characteristics, as described above, involves establishing a logical system of categories -- a classification system of related domains, sub-domains, sub-sub-domains, etc. There are a number of problems associated with such an effort.

First is the problem of defining terms. The literature refers to inventories, taxonomies, and classifications, in describing the outcomes of such analyses. The term "inventory" refers to a system of categories without any logical structure (Travers 1980, p.6). Inventories of a store's merchandise listed alphabetically in books, or on computers, is an example of an inventory. No useful relationship in the qualities of the items is suggested by the alphabetical order. To raise a system of categories above the inventory level, it must provide a valid basis for the organization of knowledge about such terms. The measurement of unrelated inventories of child-development characteristics is rejected in this project as wasteful and harmful. Such measurement would be, only, of speculative value and, therefore, could not justify the use of resources in being developed or implemented. It could also invite unwarranted and damaging interpretations.

Bloom (1956) defines a "taxonomy" as being "so constructed that the order of terms must correspond to some 'real' order among the phenomena represented by the terms . . . a taxonomy must be validated by demonstrating its consistency with the theoretical views in research findings of the field it attempts to order" (Bloom 1956, p. 17). Travers (1980, pp. 6, 17) describes "taxonomy" as "quite a vague term, with a respectable Greek background, which makes it an excellent cloak for ignorance". Alfred North Whitehead is credited with commenting that taxonomies are the death of science . . . .(They) can crystalize thinking along certain lines and inhibit exploration along different lines".

From the above, we infer that taxonomies are, or should be, scientifically determined orders. Where they cannot be science-based, there is a danger of their being counter productive. This project is not attempting to define a scientific order, i.e., a taxonomy, of child-development characteristics.

In contrast, Bloom allows that "a classification system may have many arbitrary elements....(It) may be validated by reference to the criteria of communicability, usefulness, and suggestiveness" (Bloom, 1956, p.17). Travers refers to a classification as "a system of categories in which the categories have some logical structure" (Travers, 1980).

This project proposes that the adoption of a logical classification system of characteristics, though somewhat arbitrary, is a minimum prerequisite for the systematic evaluation of Head Start program effects and that the criteria of communicability, usefulness and suggestiveness of a unique quality, as suggested by Bloom, are appropriate for the selection of such a classification system. Accordingly, there is no requirement that categories of characteristics (i.e., the domains and sub-domain constructs) be orthogonal. They may overlap as long as they are useful in addressing some unique quality.

Beyond this minimum, this project recommends that test developers be required to investigate and establish hierarchical relationships among characteristics, wherever possible, with assurances that interpretations implicit in such hierarchies are only applied to children who have been confirmed statistically as developing along the path defined by the particular hierarchy. Wherever necessary and possible, alternative hierarchies should be identified to allow for the possible various paths being followed by different groups of Head Start children. Hierarchies, as previously explained, provide a logical structure for relating children's status at the beginning and end of a program. Hierarchies also provide the most efficient framework for making a correct determination of status with a minimum number of decisions (Travers 1980, p.7).

It should be noted that some psychologists have suggested hierarchial theories of development in various domains that have the sensitivity required to measure program effects. It is recommended that measures-development contractors consider including among the characteristics selected for measurement those that address key differentiators of development. For example, it would seem particularly appropriate that characteristics involving seriation, class inclusion, and conservation, as defined by

Piaget, be included in the cognitive domain. Similarly, characteristics selected in accordance with other theories may also be appropriate to be included in the cognitive and other domains. Characteristics here proposed for possible measurement, and those subsequently used by measures-development contractors, will be reviewed by the National Panel and other scholars to ensure that those identified as important in accordance with contemporary theory are not being neglected.

A tentative definition of domains, involving a classification system developed in accordance with the recommendations of this section, has been developed for this project. It encompasses all of the characteristics identified in the input workshops and is expandable to accommodate additional characteristics newly identified by scholars or others. This system is described in detail in the next chapter.

Criterion, Background and Other Variables

To confirm the various hypotheses made during the design phases of this project, and to identify critical entry characteristics of children that will be required for analysis, data on other than child-outcome characteristics will be required, at least during the pilot and early field-testing stages of the project. To obtain data on criterion variables to validate the characteristics associated with each construct, questionnaires and rating scales for administration to parents and teaching staff should be developed. Also, data on background variables that may be critical entry characteristics can be effectively and inexpensively provided through questionnaires, interviews and program records.

It is recommended, therefore, that the requirements for development of such measures be included in the specifications for each of the measures-development contracts. As similar information may be required in the various development efforts, the coordination contractor should assume responsibility for reviewing individual contractor requirements to avoid an unnecessary duplication of effort.

It is again recommended that the data obtained on these variables, and from the child outcome measures to be developed, be maintained as a suitable data base for the evaluation of longitudinal relationships. Furthermore, the usefulness of such a base would be very much enhanced by the inclusion of data on program effects on families, on their communities, and on the institutions serving Head Start children and families.

IV

CHARACTERISTICS TO BE MEASURED

The characteristics to be measured are defined in terms of the specific indicators of desired development identified in the input workshops and the structure developed as described below. The specific characteristics are listed in the appendices attached to this report, classified by various categories in the structure. The development of measures and, therefore, the assignment of specific characteristics to a specific measurement development effort, will also be organized by the conceptual structure here used.

Approach to Defining Characteristics

The approach began with the definition of the characteristics to be measured through analysis of the behaviors and status characteristics identified by Head Start parents, Head Start staff, and public school teachers (K-2) as indicators of desired program effects in a series of input workshops. The procedures through which these indicators were produced are described in Chapter II on methodology.

Bias and diversity issues, as described in Chapter III of this report, dictated selection of a sample population for the workshops which would produce the complete range of desired characteristics. Accordingly, a systematic attempt was made to select a diverse population from which to produce a wide variety of views. Therefore, the intent was to make the sample inclusive of the various Head Start populations, but not necessarily representative of the total population.

Approximately 13,000 statements were produced through the input workshops, all offered as evidence of the attainment of competency in 26 different areas of child development (see p. 31). These areas were selected somewhat deductively through procedures involving analysis of Head Start goals and through the recommendations of scholars with a wide range of views on child development (see p. 28). It then became necessary to conceptualize an organization for this universe of desired characteristics to serve as a logical structure for the characteristics to be included in the measurement system. This was done through processes using the following criteria:

1. All of the characteristics identified as valuable in the input workshops must be included, except those outside the scope of Head Start concerns (e.g., religious beliefs).
2. Constructs representing groups of related characteristics must be identified to focus on meaningful, validated relationships among characteristics.
3. Constructs should communicate particular information on priority concerns of Head Start parents, staff and administration.
4. Each construct should represent a somewhat unique and independent set of characteristics.
5. Within each construct there should be characteristics that assure sensitivity to program effects on children 3 to 7 years of age.

A structure involving constructs that define related groups of characteristics at various levels of generalization was selected as appropriate to the needs of this project. Social Competence had been identified in the Head Start goals as the overall construct. The term "Domains" was selected to represent the first level of subconstructs which differentiate generally understood broad areas of human development, such as health, social-emotional, cognitive, etc. The term "Dimensions" was selected to designate what are hypothesized to be related sub-groups of characteristics within each Domain. These terms, "domains" and "dimensions", are also used in the literature and in this paper to refer generally to areas of child development not so specifically defined by the classification system that was adopted. Hereafter in this paper, where they are used to refer to such an area not particularly defined by this project's proposed classification system, the terms will be spelled with a lower case "d"; where "Domains" or "Dimensions" are used to refer to one or more of the specific categories under this project's classification system they will be spelled with a capital "D".

It should be noted that this conceptualization is not based on theory. It has only been selected as appropriate to the criteria guiding the process of defining characteristics to be measured in this project. Also, domains and dimensions under this structure need not be orthogonal and, as they are recommended to be used, are not. Finally, in defining a structure of characteristics consisting of constructs with somewhat defined boundaries, to facilitate communication of discrete information as required for this project, it is not implied that children's behavior takes place in such isolated or discrete segments.

Classifying Individual Characteristics

Statements from the first day of the input workshops that were similar were combined in preparing the rating-questionnaires used on the second day. An Aggregate of 4,896 characteristics were evaluated through the use of second-day questionnaires administered in the seven input workshops. Subsequently, the Mediax staff identified 1,516 somewhat unique characteristics among all of those evaluated. The original 13,000 statements are recorded in previous publications (Mediax Associates, 1978); the reduced number of characteristics are listed in the appendices to this paper, grouped by their various properties as identified by the classification system adopted[1], and referenced to the workshop questionnaries in which they appeared.

It is from the analysis of the properties of the specific characteristics identified in the input-workshops that the sub-constructs and constructs have been inductively derived. Thus, the proposed structure for measuring Head Start program effects on children stems directly from the indicators of desired development identified by the Head Start community. Two interdependent systems of classifying are described below. The first organizes behavior and development according to the process and content of the individual characteristics; the second is by four "types" of characteristic.

---

[1] See listing for Appendices B, C and D in "Contents" (p. vi).

Process-Content Categories

Each characteristic in the reduced aggregation describes a mental or physical process or status which is the basis for differentiating the various domains that have been previously defined for study by scholars. These include, for example, identifying, understanding and problem solving in what is widely accepted as the cognitive domain; enjoying, fearing and empathizing, all emotions that are associated with the affective; willing and initiating, tendencies and predispositions which some define as the conative; receiving corrective and preventative treatments which are accepted as contributing to health status; developing physically and the motor skills which are all associated with the condition "physical fitness"; and analyzing situations and planning and carrying out strategies which some call executive skills, not necessarily part of any domain. These properties are described under a "process" category in the classification system. A listing of all 1,516 characteristics grouped by process is in Appendix B. Table V lists the number and percentage of characteristics of each process category.

In addition, most of the characteristics refer to a specific content or subject which is addressed by the process. Children, for example, are to be able to identify (process) -- sounds, or customs, or school rules, or numbers, or music, or different plant species. Similarly children may enjoy -- tastes, peers, reading, art or going to the doctor. These properties of the characteristics are described under a "content" category in this classification system. Appendix C is a listing of all characteristics grouped by their specific content or, for those few in which a specific content was not prescribed, as "unspecified"; Table VI lists the number and percentage of characteristics of each content category.

Twenty-seven process categories and 31 content categories were used to classify characteristics. However, not all process-content combinations are represented; only 225 of the 810 possible combinations (all numbers less "unspecified" and "non-specific" categories) are prescribed by the characteristics identified in the workshops. Figure 2 summarizes the process-content categories in which the universe of desired characteristics identified in the input-workshops has been classified.

It should be noted that specific categorizations of characteristics may be appropriately challenged. The procedure of categorizing consistently involved choosing between reasonable alternatives; the language used in describing desired behaviors has not evolved taxonomically. On the contrary, the generally primitive state of the science of human development was apparent in the imprecision of the language. To "know", for example, can involve simple identification, or great understanding of cause and effect principles, or reflect an attitude of faith. Ambiguity is more the rule than the exception. During the process of interpreting and organizing characteristics, these ambiguities were resolved according to rules summarized in Appendix A. The general principles applied were to interpret all statements in terms of the context of Head Start and of the specific input workshops. Also, where a higher or lower level of development could be inferred, the lower level was assumed to be intended.

Although the interpretation and placement of individual characteristics may, therefore, be subject to challenge, it is believed that they may be appropriately used in defining characteristics and hypothesized constructs to be measured in this project. Precise definitions and validation of intended meanings is required and, of necessity, must be undertaken during the measures development phase. It is not anticipated that these will result in significant changes in what has been defined.

TABLE V

Number and Percent of Characteristics of Each Process Type

Code Number	Process Type	Number of Characteristics	Percent of Total (1516)
101.	Remembering	14	.92
102.	Discriminating	12	.79
103.	Sorting/Matching	26	1.72
104.	Identifying	294	19.39
105.	Recognizing Patterns & Sequences	92	6.07
106.	Understanding	138	9.10
107.	Problem Solving	7	0.46
108.	Critical Thinking	8	.53
109.	Creating; Inventing	18	1.19
199.	Unspecified	12	.79
201.	Attending	47	3.10
202.	Fearing/ Not Fearing	156	10.30
203.	Enjoying; Liking	64	4.22
204.	Expecting Success	16	1.06
205.	Believing	24	1.58
206.	Refraining; Respecting	215	14.18
207.	Willing, Wanting to Do	77	5.08
208.	Being Curious	15	.99
209.	Initiating	28	1.85
210.	Empathizing; Helping	70	4.62
301.	Checkups	6	.40
302.	Physical Status	26	1.72
303.	Physical Skills	100	6.60
401.	Analyzing	9	.59
402.	Planning	6	.40
403.	Managing	15	.99
404.	Coping with Problems	11	.73
405.	Coping with Feelings	10	.66

TABLE VI

Number and Percent of Characteristics of Each Content Type

Code Number	Content Type	Number of Charac-teristics	Percent of Total (1516)
01.	Visual	43	2.84
02.	Auditory	9	.59
03.	Haptic	44	2.90
04.	Smell	6	.40
05.	Taste	4	.26
06.	Senses, non-specific	12	.79
07.	Self	126	8.31
08.	Family and own culture	98	6.46
09.	Peers	153	10.09
10.	Adults	95	6.27
11.	School	80	5.28
12.	Community and other cultures	54	3.56
13.	Oral language	78	5.15
14.	Written language	92	6.07
15.	Second or other language	20	1.32
16.	Numbers	54	3.56
17.	Weights and measures	4	.26
18.	Money	8	.53
19.	Time	14	.92
20.	Body movement; sports; dance	65	4.29
21.	Music	13	.86
22.	Visual arts	21	1.39
23.	Dramatic arts	17	1.12
24.	Poetry and literature	14	.92
25.	Earth/universe; physical science	21	1.39
26.	Plants, animals, ecology, human physiology	66	4.35
27.	Health, nutrition, dental health	191	12.60
28.	Business; manufacturing; economics	18	1.19
29.	Building and construction	8	.53
30.	Transportation	12	.79
31.	Society; social order	33	2.18
99.	Unspecified	43	2.84

Figure 2

Process and Content of Characteristics Identified in All Domains unvalidated

CONTENT

PROCESS	VISUAL 01	AUDITORY 02	HAPTIC 03	SMELL 04	TASTE 05	NON-SPECIFIC 06	SELF 07	FAMILY 08	PEERS 09	ADULTS 10	SCHOOL 11	COMMUNITY 12	ORAL LANGUAGE 13	WRITTEN LANGUAGE 14	2ND LANGUAGE 15	NUMBERS 16	WTS. & MEASURES 17	MONEY 18	TIME 19	BODY MOVEMENT 20	MUSIC 21	VISUAL ARTS 22	DRAMATIC ARTS 23	POETRY & LITERATURE 24	EARTH/UNIVERSE 25	PLANTS, ANIMALS 26	HEALTH, NUTRITION 27	BUSINESS & MANUF. 28	BUILDING 29	TRANSPORTATION 30	SOCIETY 31	UNSPECIFIED 99
REMEMBERING 101	●	●									●		●	●	●	●																
DISCRIMINATING 102	●		●	●			●	●	●					●			●			●												
SORTING 103	●	●	●			●							●	●		●																
IDENTIFYING 104	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●			●	●	●	●	●	●	●	●	●
RECOGNIZING SEQUENCES 105	●						●		●		●		●	●		●		●	●						●	●	●	●	●	●	●	●
UNDERSTANDING 106	●					●	●	●	●	●	●	●	●	●	●	●	●	●					●		●	●	●	●	●	●	●	
PROBLEM SOLVING 107									●		●																					●
CRITICAL THINKING 108							●			●				●																		●
CREATING, INVENTING 109							●							●						●	●	●	●	●								●
UNSPECIFIED 199		●									●					●																
ATTENDING 201							●	●	●	●	●	●	●	●							●	●	●	●	●	●	●					
FEELING FREE, NOT FEARING 202						●	●	●	●	●	●	●	●	●	●	●				●	●	●	●			●	●					●
ENJOYING, LIKING 203							●	●	●	●	●	●		●						●	●	●	●	●		●	●					
EXPECTING SUCCESS 204							●			●			●	●	●					●												
BELIEVING, VALUING 205							●	●	●			●									●	●				●					●	
REFRAINING, RESPECTING 206							●	●	●	●	●	●	●		●					●						●	●				●	
WILLING TO DO 207							●	●	●	●	●			●						●							●					●
BEING CURIOUS 208							●			●	●	●		●		●									●	●		●				●
INITIATING 209							●		●	●	●			●													●					
EMPATHIZING 210								●	●	●	●	●														●						
CHECKUPS 301																											●					
PHYSICAL STATUS 302	●	●				●	●																				●					
PHYSICAL SKILLS 303			●											●						●	●	●										
ANALYZING 401							●	●																								●
PLANNING 402							●		●																							●
MANAGING 403									●	●	●		●						●				●									●
COPING/PROBLEMS 404								●	●	●		●																				●
COPING/FEELINGS 405							●		●																							●

Within specific cells of the process-content groupings, various sub-groupings reflect competency in different areas. Various specific characteristics under the same process-content category may contribute to competency in very different areas of development. For example, characteristics under the general process-content category of identifying family customs (104.08) may be indicators of cognitive competencies, or of social competencies or, with regard to eating and hygiene, health competencies. Also, it is necesssary to relate characteristics to the competency objectives of Head Start. The Domain-Dimension structure previously outlined was developed to further organize characteristics under areas of competency. Another categorization of the characteristics contributes to the definition of that structure.

Type Categories

It is useful to further classify characteristics by the type of their respective processes. Historically, three modes of mental functioning were distinguished: cognition (the acquisition and processing of knowledge), conation (volition or tendencies to act) and affection (feelings and emotions). In addition there are two categories of physical functioning: physiological and motor (English & English, 1958).

It is well recognized that human functioning involves interrelated internal processes. No one of the categories described above can operate independently. Cognitive development is inherently dependent on conative tendencies, influenced by the affective and must involve the physiological if not always the motor. The conative, similarly, is influenced if not determined by the combined effects of the cognitive, affective and physical; and the affective by the other three. The terms psychosomatic and psychomotor attest to the interrelationship between the physical and the others. The competency goals for children must, therefore, involve the interrelatedness of all these types.

Nevertheless, it is useful to define type categories, for the purposes of this project and Head Start evaluations, in conceptualizing and understanding a classification system, in defining characteristics to be measured and, very possibly, in analyzing the data from the use of the measures. Four types have, therefore, been adopted for further categorizing all characteristics by their process property — (1) knowledges and skills, (2) attitudes and habits, (3) physiological statuses, and (4) executive skills.

Knowledges and skills. Characteristics involving, primarily, a process of knowing, or what are commonly referred to as "thinking skills", are grouped under the type category "Knowledges and Skills". They include the specific processes of remembering, discriminating, sorting, identifying, recognizing sequences, understanding, problem solving, critical thinking, and creating-inventing.

Knowledges and skills are grouped together because they appear to be inextricably intertwined, at least as regards the measurement concerns of this project. For example, skills necessarily involve "knowing how" to do something, whether it be a motor behavior or an intellectual behavior. It is impossible to conceive a situation in which a child could be described as "skilled" without knowing how to perform an act. Similarly, the very act of acquiring knowledge involves some thinking skill, short-term memory as a minimum.

Within the knowledges and skills category, a hierarchy is hypothesized, extending from remembering at the lowest level to creating-inventing at the highest. Another dimension with hierarchical relationships implicit within the conceptualized category of "Knowledges and Skills" relationships is in the extent of content. Although these hierarchal relationships are noted as possibly useful in measuring change in specific areas of competency, no measure of a construct representing the overall catgory of "knowledges and skills" is recommended. Focusing on such a theoretical aggregation is not directly related to the development and measurement of social competence.

A set of processes related to those included under "knowledges and skills" is that involved in stages of cognitive development identified by Piaget -- conservation, seriation and class inclusion. Although none of the characteristics identified in the workshops appeared to be stated in Piagetian terms, manifestations of the attainments of these stages are included within this category. It is recommended that measures developers include in their measures specific Piagetian indicators of development so that relationships can be investigated from the perspective of Piaget's widely accepted theory.

Attitudes and habits. Attitudes and habits are similarly difficult to separate in measurement, if not in definition. Attitudes are defined as enduring, learned predispositions to feel, behave, or believe in a consistent way toward a given class of objects or events; and habits as acquired motor or mental acts that are relatively fixed, consistent, almost automatic. Thus, a habit always reflects some attitude; and an attitude always involves some habitual behavior. Both share qualities that are difficult to separate logically, and practically impossible to separate for purposes of measurement. "Attitudes and habits" are, therefore, defined in this project as one type of characteristic. They include the specific mental and emotional processes related to attending, feeling free and not fearing, enjoying, expecting success, believing, respecting, being willing, wanting, being curious, initiating, empathizing - or helping and related characteristics.

Although these processes do not have a clear hierarchical or other relationship, they all share the quality of representing an attitude which is at least conceptually distinguishable from the other types. Attitudes are uniquely susceptible to reversal as a function of perceived environment; children's states can instantaneously change from attending to being distracted from, from being unafraid to fearing, and from expecting success to expecting failure. Also, whereas more is almost always desirous in knowledges and skills, there is almost always a condition of "too much" which is undesirable in any area of attitudes and habits. These properties are important to the goals of Head Start and the objectives of this project. They present a critical challenge to the development of measures of program effects which, of necessity, must include this important type of characteristic.

Physiological statuses. A third type of characteristic, clearly distinguishable from the other two, involves metabolic and motor functions in contrast to the thinking skill and mental-emotional processes involved in "knowledges and skills" and "attitudes and habits". Although metabolic and motor functions have very different properties, they are here combined because of their common concern with physical potential and, specifically, the child's ability to physically function and endure.

Some desirable program effects on physical functioning and endurance are not easily measured through observations of children's behaviors. For example, it would not be practical to directly measure children's increased resistance to disease. In such cases, the receipt of a preventive treatment, such as innoculations, is presumed to be a desirable characteristic of children, and has been so identified in the input workshops. The specific processes included under the type category "physiological statuses" are preventative care, corrected conditions, physical status and physical skills.

Some of these processes have more or less broadly prescribed optimum conditions, e.g., have regular medical check-ups, hear clearly, and be of appropriate weight for height and body frame. Others, such as motor skills, are characterized by a continuum of development limited only by the physical maturity of the child.

Executive skills. In defining these separate types of characteristics, it also became logically necessary to allow for those desired characteristics which involved the effective integration of other types of characteristics appropriate to a particular real life situation. Examples of these identified in the input workshops are:

- realize what tasks they are competent to undertake, and finish, before undertaking them.
- make plans depending on their ability to respond effectively and the degree of danger.
- notice mood changes in others and adapt.
- work out disputes by talking about the situation without help of parents or teachers.

Each of these, presented as in a hypothetical situation, can be considered as primarily involving knowledge and thinking skills, such as in problem solving. In real life situations, however, each also involves attitudes and habits, such as attention, fear, empathy, expectations of success, and initiative and, often, also one's own physiological status such as the condition of one's sensory organs, general stamina and particular motor skills. The term "executive skills" has been selected to encompass that type of characteristic whose process involves the integration and organization of other types of characteristics to devise and implement effective courses of action in specific, real-life situations.

Bronson (1974) uses the concept "executive skills" with precisely the same meaning as this project. She defines such skills as the "preschool child's functional level of competence in social interaction and in coping tasks . . .skill in recognizing or identifying cues in a situation, skill in predicting possible sequences of events and outcomes of a situation, and skill in controlling both the self and the social or material 'other' in situations in order to effectively reach goals".

Defining Domain Categories

The use of domains for ordering educational or psychological information is of recent origin. As in the history of most branches of scientific knowledge, the study began with observations and description. Wright (1967) has indicated that prior to the middle of this century, most descriptive accounts of children's development were thoroughly mixed with descriptions of their social and emotional development, peception, verbal, cognitive and motor skills. Freud was the first person to introduce the notion that stages of development are sequential. However, it was Arnold Gesell, one of the first to study children over time using scientific procedures, who was also the first to put forward four major fields of child behavior: (1) motor behavior, (2) adaptive behavior, (3) language behavior, and (4) personal-social behavior. They are the precursors of domains of child development frequently in use today.

Although any particular set of domains of child development is somewhat arbitrary, domains of child development have evolved logically and become standardized through practice. For example, the Denver Developmental Screening Test, utilizing a modification of Gesell's domains, was devised to evaluate the following aspects of the child's functioning: 1) gross motor, 2) fine motor adaptive, 3) language, and 4) personal-social (Frankenburg and Dodds, 1967). Similarly, the Memphis Developmental Screening Scale evaluates in five developmental areas: 1) personal-social, 2) gross motor, 3) fine motor, 4) language and 5) percepto-cognitive (Quick, 1972). Bloom and his colleagues arranged information about educational objectives into three domains: cognitive, affective, and psychomotor. With the publication of Taxonomy of Educational Objectives: Cognitive Domain (Bloom, 1956), the term "domain" was used more commonly to classify broad catgories of child-development characteristics. There are many other tests and curricula that use domains for the arrangement of child characteristics in an orderly system.

Head Start research and evaluation studies have delineated domains, using some variations of the social-emotional, psychomotor, and cognitive domains. Physical Development, especially as it relates to health and nutritional concerns, frequently has been used to represent still another domain (Love, Wacker & Meece, 1975; Walker, Bane & Bryk, 1973).

Two other domains of developmental outcomes frequently discussed in related literature are Aesthetic Development and Moral and Ethical Development, both of which are logically essential components of Head Start's goal of social competence. Neither is represented substantially among the characteristics identified in this project, nor are they addressed in Head Start's Performance Standards. Both are areas for which the practicality of effective measurement (in the present state of the art) is least likely. For all these reasons, they are not among the domains further considered for measurement in this project.

These prior classifications of domains of child development influenced the definitions in this system, but no one of them fully sufficed. Analysis of the characteristics identified by the input workshop participants revealed the need for a domain structure somewhat -- but not fundamentally -- different from that used in previous studies.

The characteristics concerned with perceiving, knowing and thinking competencies, including the use of language, are generally conceptualized as in the Cognitive Domain, and are similarly identified for the purposes of this project. So, also, with the characteristics related with social or emotional competencies; they are identified as in the Social-Emotional Domain. Characteristics associated with the areas of health, nutrition and psychomotor development are closely related and are grouped in the Health and Physical Development Domain.

But there is also a need to define a domain that involves more than just having cognitive, social-emotional or physical competencies, one that focuses on the ability to apply them effectively in concert, to particular situations, the competency to adapt to varying environmental contexts. They constitute what is designated as the domain of "Applied Strategies".

Applied Strategies is conceptualized as being particularly concerned with interpersonal and task competency in real life situations. It includes what Lee C. Lee describes in her report to this project as "communication competence":

> . . . a dynamic process which draws heavily on cognitive, linguistic, and social capabilities. It is the translation of these

> capabilities into functionally appropriate interpersonal strategies for use in particular situational and/or socio-cultural contexts (Lee, 1978).

A similar definition, involving varied capabilities and their application through functionally appropriate task strategies for use in particular situations, could describe the non-interpersonal or task competencies. In specific learning situations, in school, for example, such competency would manifest itself in a child's initiating and carrying out purposeful learning experiences, in displaying persistence, in completing assigned and self-initiated learning tasks, and displaying what is commonly referrred to as effective learning attitudes. Such competency would also be manifested in a person's coping strategies in a wide range of out-of-school situations.

It is apparent that the four Domains closely parallel the previously defined four types of characteristics. And, as conceptualized, a different type of characteristic more closely manifests the unique quality of competence in each of the Domains --- "knowledges and skills" in the Cognitive Domain, "attitudes and habits" in the Social-Emotional Domain, "physiological statuses" in the Domain of Health and Physical Development and "executive skills" in the Applied Strategies Domain. But defining each Domain to exclusively encompass one type of characteristic would not reflect what is generally understood as comprising competency in their respective areas.

The workshop participants appropriately identified, in varying degrees, more than one type of characteristic in all of the areas comprising each Domain. In the Cognitive Domain a very large percentage of the characteristics were knowledges and skills; only a small number (between 2 and 20 percent) of "attitudes and habits" were identified as representative of desired development. However, in the areas of the Social-Emotional Domain, much larger percentages (between 10 and 40 per cent) of those identified were of the "knowledges and skills" type. And large percentages of both "attitudes and habits" and "knowledges and skills" were identified in the areas of health. These findings and subsequent analysis have led to this project's adopting a conceptualization of the relationship between types of characteristics and the four Domains that differs to some extent from those previously developed.

The Cognitive Domain is conceptualized as embracing only knowledges and skills and only in certain content areas defined as associated with Cognitive competencies. (Knowledge of good health habits, for example, is included under the Domain of Health and Phsyical Development.) Although some people associate attitudes and habits with cognitive development, such as those attitudes that contribute to learning and reflect enjoyment of cognitive activites, there appear to be more common understanding, interest and value in including only knowledges and skills in the Cognitive Domain. It appears more useful, for example, to report through the measures of this project that children have attained a level of competency in reading and, to report separately on their attitudes toward learning (including reading), rather than trying to combine both in one construct. Also, we have provided for the integration of attitudes and cognitive competencies in real life situations, such as at school, in the Applied Strategies Domain.

In the Social-Emotional Domain, however, it does not seem appropriate to include only one type of characteristic, i.e., attitudes and habits. Input workshop participants identified, for example, "understanding the meaning of facial expressions" as an indication of desired development in the areas of "sensitivity to feelings of others" and "peer relations", "knowing how their actions affect others and their importance in the family-school community" as a component of desired development in the area of "self concept", and "being capable of carrying out duties and making judgments on the basis of one's own values" as important to the development of "Independence". Accordingly, the Social-Emotional domain has been defined to include two types of characteristics -- attitudes and habits and knowledges and skills.

In the areas of the Domain of Health and Physical Development, participants included three types of characteristics -- knowledges and skills, attitudes and habits and physiological statuses. In the area of dental health, for example, it has been identified as important for a child to know what causes cavities, to develop attitudes and habits (eating and tooth-brushing), and, also, to have their cavities filled. This pattern also seems to apply to other areas of Health and Physical development. Therefor, the three types of characteristics are included under the Domain of Health and Physical Development.

Finally, the Applied Strategies Domain has been defined as integrating all relevant competencies in the other three Domains and involving executive skills to the effective handling of real-life situations. It, therefore, follows that it must also include all four types of characteristics.

The hypothesized relationship between types of characteristics and the four Domains is summarized in Figure 3. It should be noted that one particular type of characteristic appears to be more influential for each of the several Domains, these types are underlined in the row for each Domain in Figure 3. Thus, knowledges and skills are conceptualized as critical in the Cognitive Domain, attitudes and habits in the Social-Emotional Domain, physiological statuses in the Health and Physical Domain, and executive skills in Applied Strategies. There are, of course, functional interrelations among the four types of characteristics in all Domains; but the content of each Domain -- and hence the kinds of data to be collected for measurement -- includes a particular set of characteristic types that differ from the sets included in the other Domains.

Figure 3

Types of Developmental Characteristics to be Measured in the Four Domains

	Type of Characteristic			
Domain	Knowledges and Skills	Attitudes and Habits	Physiological Statuses	Executive Skills
Applied Strategies	*	*	*	*
Health & Physical	*	*	*	
Social Emotional	*	*		
Cognitive	*			

Within this structure of Domain constructs, specific subconstructs and their respective related characteristics must be identified to focus on meaningful relationships and areas of particular concern to the Head Start community. These are further defined in the following discussion of each of the Domains.

Cognitive Development

Cognition is the generic term for processes by which an individual obtains and uses knowledge. Bloom (1956) and his colleagues, in developing a taxonomy of educational objectives, defined the cognitive domain to include those objectives that deal with the recognition or recall of knowledge and the development of intellectual abilities and skills (here termed information-processing or thinking skills). So conceived, cognitive development clearly includes language development, which is often treated as a separate domain. The rationale for such inclusion is stated by Cazden (1971, p. 392):

> While we evaluate verbal behavior, we are almost always interested in the intellectual development (or cognitive structure) that such behavior presumably expresses or indicates . . . concern for (and evaluation of) specific language objectives in early childhood education must be combined with concern for (and evaluation of) cognitive development.

Cognitive development has been investigated from many points of view, two of which are especially relevant to this project. Some students of cognition focus mainly on the processes within the child through which knowledge is obtained and used, whereas other investigators attend mainly to the child's development of cognitive skills and knowledges that are demanded by the society. Both emphases are reflected in this project's selection of characteristics to be measured in the Cognitive Domain. These, in effect, specifically define the Domain for this project.

The characteristics selected for inclusion in the Cognitive Domain are listed on pages 1 to 22 in Appendix D. They are summarized by their process-content properties in Figure 4.

The measurement of cognitive development has a long and controversial history, much of which is still relevant to the issues facing this project. J. McVicker Hunt has prepared a summary of this history; it is included in Appendix H . Many of the concerns for this project emanate from the abuses, problems and misunderstandings experienced with the previous attempts at measurement that are described in this summary. The procedures used in this project to define domains for measurements and the recommendations for the kinds of measures to be developed, are intended to address these concerns.

Further analysis and classification of the characteristics included in the Cognitive Domain is required to identify sub-constructs, here called Dimensions, to be specifically addressed.

Many attempts have been made to classify dimensions of cognitive development and to order them in a way that reveals some of their essential characteristics as well as their inter-relationships. Attempts have also been made to find a psychological theory which would provide a logical and sound basis for ordering dimensions in the cognitive domain. However, it appears that no single theory of personality, learning

Figure 4

Process and Content of Characteristics Identified in the COGNITIVE Domain unvalidated

CONTENT

PROCESS	VISUAL 01	AUDITORY 02	HAPTIC 03	SMELL 04	TASTE 05	NON-SPECIFIC 06	SELF 07	FAMILY 08	PEERS 09	ADULTS 10	SCHOOL 11	COMMUNITY 12	ORAL LANGUAGE 13	WRITTEN LANGUAGE 14	2ND LANGUAGE 15	NUMBERS 16	WTS. & MEASURES 17	MONEY 18	TIME 19	BODY MOVEMENT 20	MUSIC 21	VISUAL ARTS 22	DRAMATIC ARTS 23	POETRY & LITERATURE 24	EARTH/UNIVERSE 25	PLANTS, ANIMALS 26	HEALTH, NUTRITION 27	BUSINESS & MANUF. 28	BUILDING 29	TRANSPORTATION 30	SOCIETY 31	UNSPECIFIED 99
REMEMBERING 101	●	●									●		●	●	●	●																
DISCRIMINATING 102	●		●	●			●	●						●			●															
SORTING 103	●	●	●			●							●	●		●																
IDENTIFYING 104	●	●	●	●	●	●	●	●		●	●	●	●	●	●	●	●	●	●	●	●			●	●	●		●	●	●	●	●
RECOGNIZING SEQUENCES 105	●										●		●	●		●		●	●						●	●	●	●	●	●	●	●
UNDERSTANDING 106	●					●	●	●	●	●		●	●	●	●	●	●	●					●		●	●		●	●	●	●	
PROBLEM SOLVING 107																																
CRITICAL THINKING 108							●							●																		
CREATING, INVENTING 109														●										●								
UNSPECIFIED 199		●														●																
ATTENDING 201																																
FEELING FREE, NOT FEARING 202																																
ENJOYING, LIKING 203																																
EXPECTING SUCCESS 204																																
BELIEVING, VALUING 205																																
REFRAINING, RESPECTING 206																																
WILLING TO DO 207																																
BEING CURIOUS 208																																
INITIATING 209																																
EMPATHIZING 210																																
CHECKUPS 301																																
PHYSICAL STATUS 302																																
PHYSICAL SKILLS 303																																
ANALYZING 401																																
PLANNING 402																																
MANAGING 403																																
COPING/PROBLEMS 404																																
COPING/FEELINGS 405																																

or child development is available to account for the varieties of behavior represented in the cognitive domain. Hilgard (1948) has concluded that each theory of learning accounts for some phenomena very well but is less adequate in accounting for others. Over two decades ago, Bloom (1956, p. 17) stated that "what is needed is a larger synthetic theory of learning than at present seems to be available". Although Piaget has provided such an overall theory it has not yet been applied to the range of characteristics observable in Head Start Programs and communities and which are required for this project.[2]

A pragmatic determination of the cognitive abilities to be measured in this project can be approached through analysis of the characteristics identified in the input workshops. On the basis of the criteria previously established (p. 44), the following six dimensions are recommended as sub-constructs for measurement in the Cognitive Domain.[3]

51. Perception
52. Language
53. Reading
54. Math concepts
55. Nature and Science Concepts
56. Social Organization Concepts

It should be noted again that, although no affective characteristics are identified for measurement under the Cognitive Domain, they are recognized as inseparably intertwined with children's development of cognitive skills in real life.

Perception. Cognitive skills begin to develop with the processing of sensory inputs. These processes, the thinking skills that are the basis of all cognitive function and, in particular, development in the Cognitive Domain, develop themselves through the processing of sensory inputs. The Perception Dimension is recommended to measure children's level of competency in developing basic knowledge and skills through sensory inputs.

Perception is also recommended for measurement because it is a prerequisite to development in almost all other areas in all Domains. Data on its level of development can, therefore, be a critical factor in analyzing all program effects on children.

The characteristics included under the Perception Dimension are listed on pages 1 - 3 of Appendix D. They are summarized by their process-content properties in Figure 5. The identified characteristics involve the lower level thinking skills with all five sensory inputs -- visual, auditory, haptic, smell and taste. The more advanced perceptual skills are included with the other dimensions they more specifically address, for example, those visual skills that are specifically involved in reading and math competency; those auditory that are involved in language; those haptic with fine motor; and, similarly, smell and taste with nutrition.

---

[2] For further comment on the applicability of Piagetian theory see Comments of J. McVicker Hunt in Appendix H.

[3] For explanation of assigned code numbers see Appendix A.

Figure 5

Process and Content of Characteristics Identified in Dimension 51: PERCEPTION unvalidated

CONTENT

PROCESS	01 VISUAL	02 AUDITORY	03 HAPTIC	04 SMELL	05 TASTE	06 NON-SPECIFIC	07 SELF	08 FAMILY	09 PEERS	10 ADULTS	11 SCHOOL	12 COMMUNITY	13 ORAL LANGUAGE	14 WRITTEN LANGUAGE	15 2ND LANGUAGE	16 NUMBERS	17 WTS. & MEASURES	18 MONEY	19 TIME	20 BODY MOVEMENT	21 MUSIC	22 VISUAL ARTS	23 DRAMATIC ARTS	24 POETRY & LITERATURE	25 EARTH/UNIVERSE	26 PLANTS, ANIMALS	27 HEALTH, NUTRITION	28 BUSINESS & MANUF.	29 BUILDING	30 TRANSPORTATION	31 SOCIETY	99 UNSPECIFIED
REMEMBERING 101	●	●									●																					
DISCRIMINATING 102	●		●	●			●	●																								
SORTING 103	●	●	●			●										●																
IDENTIFYING 104	●	●	●	●	●	●							●												●	●						●
RECOGNIZING SEQUENCES 105	●																															●
UNDERSTANDING 106	●					●																										
PROBLEM SOLVING 107																																
CRITICAL THINKING 108																																
CREATING, INVENTING 109																																
UNSPECIFIED 199																																
ATTENDING 201																																
FEELING FREE, NOT FEARING 202																																
ENJOYING, LIKING 203																																
EXPECTING SUCCESS 204																																
BELIEVING, VALUING 205																																
REFRAINING, RESPECTING 206																																
WILLING TO DO 207																																
BEING CURIOUS 208																																
INITIATING 209																																
EMPATHIZING 210																																
CHECKUPS 301																																
PHYSICAL STATUS 302																																
PHYSICAL SKILLS 303																																
ANALYZING 401																																
PLANNING 402																																
MANAGING 403																																
COPING/PROBLEMS 404																																
COPING/FEELINGS 405																																

Language. The Language Dimension addresses listening and speaking competencies, the essentials for early communication. Language has been selected because it is one of the most basic skills and, as such, is a prerequisite to further development in many other areas. It is also one emphasized in the Head Start Basic Educational Skills demonstration program.

Figure 6 summarizes the process-content properties of the characteristics hypothesized as indicators of desired language development; on pages 4 - 8 of Appendix D the specific characteristics are listed.

Reading. Reading, includes both the essential pre-reading skills, appropriate for 3-year-old Head Start children, and the more advanced skills appropriate for advanced 7 year olds. Also the skills included under this dimension include both reading and writing. They are recommended because they are also a high priority concern of all of Head Start and the recent Basic Educational Skills initiative.

The specific characteristics included under Reading are listed on pages 9 - 11 of Appendix D. They are summarized by their process-content categories in Figure 7.

Math Concepts. A Math Concepts Dimension is selected for measurement because it, also, is a priority concern of all Head Start and the Basic Educational Skills program. It includes pre-math perceptual skills, number concepts, basic computational skills, and their application to simple everyday problems.

Specific characteristics identified under Math Concepts are listed on pages 12 - 15 of Appendix D. They are summarized by their process-content categories in Figure 8.

Nature and Science. The previous four Dimensions in the Cognitive Domain have been concerned with the development of basic cognitive skills -- peception, language, reading and math -- and include the basic thinking skills involved in each. In this and the next Dimension, the primary concern is for the specific content of the knowledge; the basic knowledges and thinking skills processes -- remembering, discriminating and sorting -- are, therefore, not addressed. The differences in the nature of the skills in this Dimension, as compared with the previously recommended Dimensions for the Cognitive Domain, are evident from a comparison of their respective process-content summaries.

This Dimension, Nature and Science has been selected to measure children's basic knowledge in such areas as weather, seasons, day and night, relationships of rain, snow and ice, natural resources, geology, the plant and animal world, ecology, the unique nature and role of human beings, the needs of human beings, and the products of science and technology. Specific characteristics recommended to be included under the Nature and Science Dimension are listed on pages 16 - 19 of Appendix D. A summary of their process-content properties is presented in Figure 9.

Social Organization. The nature of the recommended Dimension Social Organization is similar to Nature and Science except for the subject matter. Here the concern is for the child's knowledges and skills of societal institutions which affect so much of one's interpersonal relationships. This Dimension includes concern for the particular problems confronted by children of minority backgrounds. Not only must they learn to deal effectively with the customs and structures of their own culture, but also they must learn about the "foreign", and therefore more difficult, institutions of the mainstream society.

Figure 6

Process and Content of Characteristics Identified in Dimension 52: LANGUAGE unvalidated

CONTENT

PROCESS	01 VISUAL	02 AUDITORY	03 HAPTIC	04 SMELL	05 TASTE	06 NON-SPECIFIC	07 SELF	08 FAMILY	09 PEERS	10 ADULTS	11 SCHOOL	12 COMMUNITY	13 ORAL LANGUAGE	14 WRITTEN LANGUAGE	15 2ND LANGUAGE	16 NUMBERS	17 WTS. & MEASURES	18 MONEY	19 TIME	20 BODY MOVEMENT	21 MUSIC	22 VISUAL ARTS	23 DRAMATIC ARTS	24 POETRY & LITERATURE	25 EARTH/UNIVERSE	26 PLANTS, ANIMALS	27 HEALTH, NUTRITION	28 BUSINESS & MANUF.	29 BUILDING	30 TRANSPORTATION	31 SOCIETY	99 UNSPECIFIED
REMEMBERING 101	●												●																			
DISCRIMINATING 102																																
SORTING 103													●																			
IDENTIFYING 104	●		●	●	●	●	●	●			●		●		●						●											
RECOGNIZING SEQUENCES 105	●										●		●																			●
UNDERSTANDING 106													●										●									
PROBLEM SOLVING 107																																
CRITICAL THINKING 108																																
CREATING, INVENTING 109																								●								
UNSPECIFIED 199		●																														
ATTENDING 201																																
FEELING FREE, NOT FEARING 202																																
ENJOYING, LIKING 203																																
EXPECTING SUCCESS 204																																
BELIEVING, VALUING 205																																
REFRAINING, RESPECTING 206																																
WILLING TO DO 207																																
BEING CURIOUS 208																																
INITIATING 209																																
EMPATHIZING 210																																
CHECKUPS 301																																
PHYSICAL STATUS 302																																
PHYSICAL SKILLS 303																																
ANALYZING 401																																
PLANNING 402																																
MANAGING 403																																
COPING/PROBLEMS 404																																
COPING/FEELINGS 405																																

Figure 7

Process and Content of Characteristics Identified in Dimension 53: READING unvalidated

CONTENT

PROCESS	01 VISUAL	02 AUDITORY	03 HAPTIC	04 SMELL	05 TASTE	06 NON-SPECIFIC	07 SELF	08 FAMILY	09 PEERS	10 ADULTS	11 SCHOOL	12 COMMUNITY	13 ORAL LANGUAGE	14 WRITTEN LANGUAGE	15 2ND LANGUAGE	16 NUMBERS	17 WTS. & MEASURES	18 MONEY	19 TIME	20 BODY MOVEMENT	21 MUSIC	22 VISUAL ARTS	23 DRAMATIC ARTS	24 POETRY & LITERATURE	25 EARTH/UNIVERSE	26 PLANTS, ANIMALS	27 HEALTH, NUTRITION	28 BUSINESS & MANUF.	29 BUILDING	30 TRANSPORTATION	31 SOCIETY	99 UNSPECIFIED
REMEMBERING 101														●																		
DISCRIMINATING 102														●																		
SORTING 103													●	●																		
IDENTIFYING 104													●	●	●									●								
RECOGNIZING SEQUENCES 105														●																		
UNDERSTANDING 106														●	●								●									
PROBLEM SOLVING 107																																
CRITICAL THINKING 108							●							●																		
CREATING, INVENTING 109														●										●								
UNSPECIFIED 199																																
ATTENDING 201																																
FEELING FREE, NOT FEARING 202																																
ENJOYING, LIKING 203																																
EXPECTING SUCCESS 204																																
BELIEVING, VALUING 205																																
REFRAINING, RESPECTING 206																																
WILLING TO DO 207																																
BEING CURIOUS 208																																
INITIATING 209																																
EMPATHIZING 210																																
CHECKUPS 301																																
PHYSICAL STATUS 302																																
PHYSICAL SKILLS 303																																
ANALYZING 401																																
PLANNING 402																																
MANAGING 403																																
COPING/PROBLEMS 404																																
COPING/FEELINGS 405																																

Figure 8

Process and Content of Characteristics Identified in Dimension 54: MATH CONCEPTS
unvalidated

CONTENT

PROCESS	VISUAL 01	AUDITORY 02	HAPTIC 03	SMELL 04	TASTE 05	NON-SPECIFIC 06	SELF 07	FAMILY 08	PEERS 09	ADULTS 10	SCHOOL 11	COMMUNITY 12	ORAL LANGUAGE 13	WRITTEN LANGUAGE 14	2ND LANGUAGE 15	NUMBERS 16	WTS. & MEASURES 17	MONEY 18	TIME 19	BODY MOVEMENT 20	MUSIC 21	VISUAL ARTS 22	DRAMATIC ARTS 23	POETRY & LITERATURE 24	EARTH/UNIVERSE 25	PLANTS, ANIMALS 26	HEALTH, NUTRITION 27	BUSINESS & MANUF. 28	BUILDING 29	TRANSPORTATION 30	SOCIETY 31	UNSPECIFIED 99
REMEMBERING 101															●	●																
DISCRIMINATING 102																	●															
SORTING 103	●		●													●																
IDENTIFYING 104	●						●									●	●	●	●									●				
RECOGNIZING SEQUENCES 105																●		●	●													
UNDERSTANDING 106																●	●	●														
PROBLEM SOLVING 107																																
CRITICAL THINKING 108																																
CREATING, INVENTING 109																																
UNSPECIFIED 199																●																
ATTENDING 201																																
FEELING FREE, NOT FEARING 202																																
ENJOYING, LIKING 203																																
EXPECTING SUCCESS 204																																
BELIEVING, VALUING 205																																
REFRAINING, RESPECTING 206																																
WILLING TO DO 207																																
BEING CURIOUS 208																																
INITIATING 209																																
EMPATHIZING 210																																
CHECKUPS 301																																
PHYSICAL STATUS 302																																
PHYSICAL SKILLS 303																																
ANALYZING 401																																
PLANNING 402																																
MANAGING 403																																
COPING/PROBLEMS 404																																
COPING/FEELINGS 405																																

Figure 9

Process And Content Of Characteristics Identified In Dimension 55: NATURE & SCIENCE unvalidated

CONTENT

PROCESS	VISUAL 01	AUDITORY 02	HAPTIC 03	SMELL 04	TASTE 05	NON-SPECIFIC 06	SELF 07	FAMILY 08	PEERS 09	ADULTS 10	SCHOOL 11	COMMUNITY 12	ORAL LANGUAGE 13	WRITTEN LANGUAGE 14	2ND LANGUAGE 15	NUMBERS 16	WTS. & MEASURES 17	MONEY 18	TIME 19	BODY MOVEMENT 20	MUSIC 21	VISUAL ARTS 22	DRAMATIC ARTS 23	POETRY & LITERATURE 24	EARTH/UNIVERSE 25	PLANTS, ANIMALS 26	HEALTH, NUTRITION 27	BUSINESS & MANUF. 28	BUILDING 29	TRANSPORTATION 30	SOCIETY 31	UNSPECIFIED 99
REMEMBERING 101																																
DISCRIMINATING 102																																
SORTING 103																																
IDENTIFYING 104																								●	●		●	●	●		
RECOGNIZING SEQUENCES 105																		●						●	●	●	●	●	●		●
UNDERSTANDING 106																								●	●		●	●	●		
PROBLEM SOLVING 107																																
CRITICAL THINKING 108																																
CREATING, INVENTING 109																																
UNSPECIFIED 199																																
ATTENDING 201																																
FEELING FREE, NOT FEARING 202																																
ENJOYING, LIKING 203																																
EXPECTING SUCCESS 204																																
BELIEVING, VALUING 205																																
REFRAINING, RESPECTING 206																																
WILLING TO DO 207																																
BEING CURIOUS 208																																
INITIATING 209																																
EMPATHIZING 210																																
CHECKUPS 301																																
PHYSICAL STATUS 302																																
PHYSICAL SKILLS 303																																
ANALYZING 401																																
PLANNING 402																																
MANAGING 403																																
COPING/PROBLEMS 404																																
COPING/FEELINGS 405																																

The specific characteristics identified with this dimension are listed on pages 20 - 22 of Appendix D. The summary of their process-content characteristics is presented in Figure 10.

Social-Emotional Development

Head Start policymakers, scholars and the participants in the input workshops of this project seem to generally agree that effective social and emotional behaviors are essential components of competence among young people. Children need to develop socially desirable relationships with other persons, together with feelings that foster such relationships and at the same time support self-realization. They also require the emotional competencies to effectively undertake and complete tasks at home, at school, and in the community. "It may well be", as Ryle (1978) suggests in a paper prepared for this project, "that the crucial factors that influence success (in school) may be the ones that are indicative of the child's underlying feelings about himself and how he relates to the rest of the world ...".

There is no such agreement, however, on the particular characteristics that constitute desirable social and emotional development; relevant literature reflects wide-ranging conceptions on this point. As Exner and London (1978) comment regarding a related analysis, there is a notable "heterogeneity of concepts studied under the rubric of dimensions of personality ... (there is) no overarching plan or theory, implicit or explicit ..." Thus, any definition of the construct and dimensions "Social-Emotional Domain" -- even more than with the other Domains -- entails selective judgments that are less guided by prevailing theory or patterns of practice and may be more concerned with optimal usefulness to the particular purposes of this project.

Jean Carew's report (1978) to the National Panel of this project makes an important distinction between "two broadly different ways of considering social development in children, social behavior development and social concept development". The differentiation is applicable also to emotional development, in the same sense that a child not only exhibits affective behaviors, but also develops knowledge and understanding about such behaviors. It is here hypothesized that concept development is functionally related to behavior development. Either may be covert or overt. Both are foci of this project. Therefore, as previously stated, measurement in the Social-Emotional Domain will be on changes in the development of attitudes and habits and knowledges and skills, although the former are hypothesized as the more critical.

Social-emotional development, however defined, is also functionally related to behavior in all other Domains; only through abstraction can it be conceived as a "separate" area of development. Nevertheless, such abstraction is warranted, for at least three reasons.

First, the "social" and "emotional" components of behaviors in other Domains are not their distinguishing characteristics. Brushing one's teeth, for example, is definitely a health-related behavior, even though it may be accompanied by positive or negative feelings toward demanding parents, or the act itself. Similarly, whatever the emotional state of a child, his or her efforts to solve an arithmetic problems, for example, clearly constitute cognitive behavior. Therefore, the weight of concern in these other Domains must be on other than their social or emotional aspects.

Figure 10

Process and Content of Characteristics Identified in Dimension 56: SOCIAL ORGANIZATION
unvalidated

CONTENT

PROCESS	VISUAL 01	AUDITORY 02	HAPTIC 03	SMELL 04	TASTE 05	NON-SPECIFIC 06	SELF 07	FAMILY 08	PEERS 09	ADULTS 10	SCHOOL 11	COMMUNITY 12	ORAL LANGUAGE 13	WRITTEN LANGUAGE 14	2ND LANGUAGE 15	NUMBERS 16	WTS. & MEASURES 17	MONEY 18	TIME 19	BODY MOVEMENT 20	MUSIC 21	VISUAL ARTS 22	DRAMATIC ARTS 23	POETRY & LITERATURE 24	EARTH/UNIVERSE 25	PLANTS, ANIMALS 26	HEALTH, NUTRITION 27	BUSINESS & MANUF. 28	BUILDING 29	TRANSPORTATION 30	SOCIETY 31	UNSPECIFIED 99
REMEMBERING 101																																
DISCRIMINATING 102																																
SORTING 103																																
IDENTIFYING 104							●	●		●		●			●			●		●				●							●	
RECOGNIZING SEQUENCES 105																															●	
UNDERSTANDING 106							●	●	●	●		●																			●	
PROBLEM SOLVING 107																																
CRITICAL THINKING 108																																
CREATING, INVENTING 109																																
UNSPECIFIED 199																																
ATTENDING 201																																
FEELING FREE, NOT FEARING 202																																
ENJOYING, LIKING 203																																
EXPECTING SUCCESS 204																																
BELIEVING, VALUING 205																																
REFRAINING, RESPECTING 206																																
WILLING TO DO 207																																
BEING CURIOUS 208																																
INITIATING 209																																
EMPATHIZING 210																																
CHECKUPS 301																																
PHYSICAL STATUS 302																																
PHYSICAL SKILLS 303																																
ANALYZING 401																																
PLANNING 402																																
MANAGING 403																																
COPING/PROBLEMS 404																																
COPING/FEELINGS 405																																

Second, there are aspects of social and emotional development which are essential to general competence at all ages. Lack of appropriate social-emotional development can be an overriding handicap, precluding effective application of attainments in other domains. A child experiencing strong emotions, or who is among strangers perceived as threatening, for example, is not likely to give his or her optimal performance in thinking.

Third, different approaches and emphases are required for nurturing the characteristics affecting social-emotional development. Therefore, the evaluation and improvement of programs necessarily involves specific measures of social-emotional characteristics.

Thus, although functionally related to the other Domains, Social-Emotional Development is different from them. Moreover, although feelings are aroused by non-social events, they are overwhelmingly influenced by factors involving interrelations of people. It is appropriate, therefore, to couple children's affective and interpersonal characteristics together, as constituting a "separate" domain of social-emotional development. This domain consists of what Minuchin, Biber, Shapiro, and Zimilies (1969, p. 8), writing about children's learning in school, characterize as:

> . . .those aspects of early development most concerned with relationships and feelings and with the ways in which a young child's experience in school affects the perception of others, the concept of self, and the expression and management of feelings.

The terms "social" and "emotional" are used in their common meanings. Children's social characteristics are those that involve the interaction of two or more people, especially interaction with peers and adults. Children's emotional behaviors are such patterns of action commonly referred to as fear, anger, interest, motivation, affection, sympathy, joy, pity, etc. They are complex behaviors in which the affective element is dominant.

There are at least two theoretical controversies in the field of personality research that have relevance for this project's definition of sub-construct Dimensions for measurement within the Social-Emotional Domain. The first issue is whether emotional reactions reflect some underlying constructs ("inner-states") that account for consistency in one's behavior. It involves the basic distinction Goodenough made three decades ago "between the 'sign' and 'sample' approaches to the interpretation of test responses. The sign approach assumes that the response may best be construed as an indirect manifestation of some underlying personality characteristic. The sample approach, on the other hand, assumes that the test behavior constitutes a subset of the actual behaviors of interest. Whereas traditional personality tests have typically taken the sign approach to interpretation, behavioral procedures approach test interpretation with the sample orientation." (Quoted in Ciminero _et al_, 1977, p. 5). Although this project does not presume to take a position on the theoretical questions implicit in this distinction, it appears that the measurement purpose of the project is better served by the behavioral approach. As Mischel (1968, p. 10) notes: "In behavioral analyses the emphasis is on what a person _does_ in situations rather than on inferences about what attributes he _has_ more globally". Overt behaviors that are accepted indicators of competency are observable and, therefore, more amenable to measurement than "inner-states" that are hypothesized to control such competent behavior.

The second issue centers around the concept "trait", a construct that usually implies generalized patterns of behavior that are consistent across contexts. As defined by Shaw and Wright (1967, p. 5), a trait is "a more or less stable and consistent disposition of the individual to respond in a certain way which differentiates him from other individuals ... traits are non-specific, generalized orientations of the individual". English and English (1958, p. 560) emphasize this more or less independence of traits from specific context, noting that a trait is a personal characteristic "which is consistently manifested, despite variations within a considerable range of circumstances." The main question relevant to this project, highlighted in Mischel's Personality and Assessment (1968), is whether understanding of one's personality is better achieved through analysis of presumably "independent" traits or, more holistically, through analysis of the dynamic interrelations of personal characteristics.

Apparently the "trait" concept, long eschewed by many investigators is experiencing somewhat of a revival. Exner and London (op. cit.) assert that much of the "newer research has been marked by a significant upswing in uniconcept dimension or trait research". In a paper prepared for this project, however, Zimiles (1978, pp. 14, 16, 17) characterizes the trait model as "obsolete", especially for measuring social and personality variables. He calls, instead, for "idiographic methods of measurement", using sustained observation by astute observers. He insists, further, that:

> The only chance for developing real measures of real children, and for approaching a valid evaluation of the impact of Head Start, is to adopt the conceptual framework of expert teachers, of people who plan and enact the programs for real children, and who, by virtue of their sustained contact and relationship with the children, can speak meaningfully about their cognitive and personality development.

This point of view has in large measure guided this project's conceptualization of the Social-Emotional Domain, and it is reflected in procedures adopted for identifying characteristics to be measured in all domains. There is no assumption here of a checklist of valued traits; rather, the focus is on observable social and emotional characteristics which are valued as indicators of competence in themselves and in the profile of their relationships. Nor is there any positing here of preconceived values in such interrelationships among characteristics. There may be such interrelationships, but it is here proposed to develop measures of program effects on individual characteristics. The methods for aggregating these characteristics into meaningful constructs of development will be determined empirically with data from what Zimiles (1978) describes as "people ... who, by virtue of their sustained contact and relationships with the children, can speak meaningfully about their personality development". It may well develop that different values held by different racial-ethnic-regional groups of such people will result in different "formulae" for aggregation for different children.

In this project, the people described by Zimiles are, specifically, the 575 Head Start parents and staff members and K-2 teachers who participated in the two-day Input Workshops conducted by this project. They identified the specific social-emotional characteristics that are the focus of the proposed measures. Their counterparts in the sample populations selected to be included in the evaluations in which the measures will be used will confirm the relevancy of these characteristics and provide the data required for their aggregation.

Although this project is concerned with program effects measurement, rather than clinical evaluations of the social-emotional characteristics of individual children, it is still important that all of the characteristics required by a holistic approach be included in our measures. Those most knowledgeable about the holistic needs of the children in particular programs can then evaluate the implications of profiles of program effects for their individual programs.

Approximately half of the characteristics identified in the input workshops are associated the Social-Emotional Domain. They are listed in pages 23 through 63 of Appendix D. (They comprise 41 of the 82 pages in the Appendix.) These characteristics are summarized by their process-content properties in Figure 11. In accordance with the criteria adopted for the selection of constructs, eight Dimensions, encompassing all these characteristics, are recommended for measurement in the Social-Emotional Domain:

61. Sensitivity to feelings of others
62. Expression of own feelings
63. Self-concept
64. Attitudes towards school
65. Independence
66. Sharing and Competing
67. Peer Relationships
68. Adult Relationships

Evident from inspection of the characteristics identified in the input workshops with desired development in the areas of these Dimensions (Appendix D) is the fact that these Dimensions do not represent unitary "traits"; rather, they are logically derived categories of specific types of desired characteristics. Illustrative are some of those included in two of the Dimensions.

Sensitivity to Feelings of Others

Know that their actions affect other people.
The meaning of sympathy.
Recognize different emotional expressions by others (e.g., facial expressions).
Display sympathy or concern for others, compassion, consideration.
Praise others who need/desire it.
Accept physical differences and handicaps of others.
Avoid ridiculing others; embarrassing others; hurting others' feelings.
Be sensitive to feelings of people of other races, nationalities, languages, etc.

Peer Relationships

Treat other children as they would like to be treated; the Golden Rule.
The concept of the rights of others.
Work and play without being overly submissive.
Work out disputes without fighting with teacher/adult help.
Make new friends.
Distinguish between intentional and accidental actions.
Help and support friends appropriately.
Want to join in group games with other children.
Be willing to listen to other children's points of view; ask what they would like to do.

Figure 11

Process and Content of Characteristics Identified in the SOCIAL-EMOTIONAL Domain unvalidated

CONTENT

PROCESS	VISUAL 01	AUDITORY 02	HAPTIC 03	SMELL 04	TASTE 05	NON-SPECIFIC 06	SELF 07	FAMILY 08	PEERS 09	ADULTS 10	SCHOOL 11	COMMUNITY 12	ORAL LANGUAGE 13	WRITTEN LANGUAGE 14	2ND LANGUAGE 15	NUMBERS 16	WTS. & MEASURES 17	MONEY 18	TIME 19	BODY MOVEMENT 20	MUSIC 21	VISUAL ARTS 22	DRAMATIC ARTS 23	POETRY & LITERATURE 24	EARTH/UNIVERSE 25	PLANTS, ANIMALS 26	HEALTH, NUTRITION 27	BUSINESS & MANUF. 28	BUILDING 29	TRANSPORTATION 30	SOCIETY 31	UNSPECIFIED 99
REMEMBERING 101											●																					
DISCRIMINATING 102								●	●											●												
SORTING 103																																
IDENTIFYING 104	●						●	●	●	●	●	●	●		●	●																
RECOGNIZING SEQUENCES 105							●		●		●		●						●													●
UNDERSTANDING 106	●						●	●	●	●	●	●	●										●								●	
PROBLEM SOLVING 107									●		●																					●
CRITICAL THINKING 108							●			●																						●
CREATING, INVENTING 109							●													●	●	●	●	●								●
UNSPECIFIED 199		●									●																					
ATTENDING 201							●	●	●	●	●	●	●	●							●	●	●	●	●	●						
FEELING FREE, NOT FEARING 202						●	●	●	●	●	●	●	●	●	●	●				●	●	●	●			●	●					●
ENJOYING, LIKING 203							●	●	●	●	●	●		●						●	●	●	●	●		●	●					
EXPECTING SUCCESS 204							●			●			●	●	●																	
BELIEVING, VALUING 205							●	●	●			●									●	●				●					●	
REFRAINING, RESPECTING 206							●	●	●	●	●	●	●		●											●	●				●	
WILLING TO DO 207							●	●	●	●	●			●													●					●
BEING CURIOUS 208							●			●	●	●		●		●									●	●		●				●
INITIATING 209							●		●	●	●			●																		
EMPATHIZING 210								●	●	●	●	●														●						
CHECKUPS 301																																
PHYSICAL STATUS 302																																
PHYSICAL SKILLS 303																																
ANALYZING 401																																
PLANNING 402																																
MANAGING 403																																
COPING/PROBLEMS 404																																
COPING/FEELINGS 405																																

Defend or get help for other children in trouble.
Make visiting children feel at home.

Under this conceptualization of eight dimensions in the Social-Emotional Domain, certain constructs of interest are omitted. J. McVicker Hunt, in a personal communication to this project, pointed out specific omissions of characteristics which imply "intentional behavior, initiative and trust". These, and other constructs of concern, are conceptualized as being of a higher order of functioning, dependent on the particular development of characteristics included within these eight Dimensions, and others in the Domains of Cognitive and Health and Physical Development, and on situational particulars. They are, therefore, more appropriately measured within the Applied Strategies Domain.

Behavior with definite goals (intentional behavior) is conceptualized as a function of knowledge in the particular area (e.g., of particular academic subjects in school of social organization at home and in the community); of problem solving ability (i.e., thinking skills); of relative expectations of success; of alternate plans of action; of the need and tolerance for change in stimulus inputs (i.e., stress factors in the medical Dimension), and of other Dimensions within the Cognitive, Social-Emotional and Health and Physical Development Domains as they relate to factors within particular situations. As such, it is recommended that they are best measured with regard to these situational factors as is conceptualized under the Applied Strategies Domain.

Similarly, initiative and motivation are conceptualized largely as a function of expectation of success (self-concept), and anticipation of reinforcements from within oneself and from the situation and, as such, are best understood and measured within the context of Applied Strategies.

And children's behavior in formal test situations such as required in school, are considered a function of Applied Strategies which includes executive test taking skills, attitudes involving motivation and concentration as well as the particular knowledge being tested.

Trust, as Hunt points out, involves a belief:

> that adults (or others being trusted) will understand what the child's intentional goals are. It shows concretely as willingness to try new and more difficult goals when they are presented by familiar adults (i.e., mothers or fathers who have made a habit of showing concern for their infants' and children's own intentional goals, rather than merely demanding what the culture demands in order to have the child serve to do credit to the adult's status as a good parent.) As Burton White (1978) has found, the "good mothers" are not those who spend the most time with their infants but those who keep an eye on them, discern what their infants are trying to do when they get frustrated, and then modify the situation so that the child can achieve his/her goal with a strong effort. This is also what Reuven Feurenstein (1976) has called" mediating the environment" for children" (Hunt, 1979).

As such, trust appears to be a higher-order construct dependent on particular relationships, available reinforcements, and risks in a situation; thus, trust is also best measured within Applied Strategies.

Relegating these higher-order constructs -- intentional behavior, motivation, trust and others, such as task orientation, persistance and effectiveness -- to a situational specific domain (i.e., Applied Strategies) does not necessarily mean that there are not consistent differences in patterns of these behaviors among children and across situations. It does imply, however, that the consistency variable is a function of the development of other specific characteristics and of the consistency in the situations themselves.

Thus, the Social-Emotional Domain consists of eight Dimensions, or subconstructs, comprising the characteristics identified in the input workshops as indicators of desired development in the various areas of the Domain. The definition of each of these Dimensions, and specific considerations for their individual selection, are provided in the following paragraphs.

Sensitivity to Feelings of Others. This Dimension is recommended to measure a child's basic ability to comprehend the feelings of others and to communicate this sensitivity to others. Both of these competencies appear prerequisite to other Dimensions of competence in the Social-Emotional Domain, for example, expression of own feelings, sharing, peer relationships and adult relationships. Although there may be significant overlap with these other Dimensions, "sensitivity to feelings of others" is recommended as a Dimension for measurement because of this seminal quality.

The characteristics include knowledges and skills in recognizing various verbal and non-verbal communications of feelings. It also includes attitudes reflecting interest, respect and concern for the feelings of others. The specific characteristics identified as comprising the Dimension of "Sensitivity to Feelings of Others" are listed on pages 23 - 25 of Appendix D. They are also summarized in accordance with their process-content properties in Figure 12.

Because of the communications nature of this Dimension, and its dependency on informal communications, it is inherently influenced by language, regional and other cultural factors.

Expression of Own Feelings. This Dimension is defined as the ability to effectively communicate one's own attitudes and feelings. As such it is considered critical to children's overall health and well being and to their present and future effectiveness in all situations.

"Expression of Own Feelings" includes the knowledges and skills to communicate one's own feelings through various verbal and non-verbal forms. It also includes, as most important, development of an attitude of openness and freedom to express (see process 202 characteristics), and development of habits of effective communication.

The characteristics comprising "Expression of Own Feelings" are listed on pages 26 - 29 in Appendix D. They are summarized by their process-content properties in Figure 13.

Self-Concept. Self concept is defined as "a personal judgment of worthiness that is expressed in the attitudes the individual holds toward himself" (Coopersmith, 1967, pp. 4-5).

Cross (1978), in a paper prepared for this project, points out a number of common misconceptions about this dimension which affect its measurement. He states that there are "two presuppositions concerning the emotional development of low income (Black/minority) children: (a) upon entrance into Head Start programs, low

Figure 12

Process and Content of Characteristics Identified in Dimension 61:
SENSITIVITY TO THE FEELINGS OF OTHERS
unvalidated

CONTENT

PROCESS	VISUAL 01	AUDITORY 02	HAPTIC 03	SMELL 04	TASTE 05	NON-SPECIFIC 06	SELF 07	FAMILY 08	PEERS 09	ADULTS 10	SCHOOL 11	COMMUNITY 12	ORAL LANGUAGE 13	WRITTEN LANGUAGE 14	2ND LANGUAGE 15	NUMBERS 16	WTS. & MEASURES 17	MONEY 18	TIME 19	BODY MOVEMENT 20	MUSIC 21	VISUAL ARTS 22	DRAMATIC ARTS 23	POETRY & LITERATURE 24	EARTH/UNIVERSE 25	PLANTS, ANIMALS 26	HEALTH, NUTRITION 27	BUSINESS & MANUF. 28	BUILDING 29	TRANSPORTATION 30	SOCIETY 31	UNSPECIFIED 99
REMEMBERING 101																																
DISCRIMINATING 102									●																							
SORTING 103																																
IDENTIFYING 104	●								●																							
RECOGNIZING SEQUENCES 105							●																									
UNDERSTANDING 106													●																			
PROBLEM SOLVING 107																																
CRITICAL THINKING 108																																
CREATING, INVENTING 109																																
UNSPECIFIED 199																																
ATTENDING 201									●	●																						
FEELING FREE, NOT FEARING 202							●	●	●	●																						
ENJOYING, LIKING 203									●																	●						
EXPECTING SUCCESS 204																																
BELIEVING, VALUING 205												●										●										
REFRAINING, RESPECTING 206									●																							
WILLING TO DO 207																											●					
BEING CURIOUS 208										●																						
INITIATING 209																																
EMPATHIZING 210								●	●	●		●																				
CHECKUPS 301																																
PHYSICAL STATUS 302																																
PHYSICAL SKILLS 303																																
ANALYZING 401																																
PLANNING 402																																
MANAGING 403																																
COPING/PROBLEMS 404																																
COPING/FEELINGS 405																																

Figure 13

Process and Content of Characteristics Identified in Dimension 62: EXPRESSION OF OWN FEELINGS unvalidated

CONTENT

PROCESS	VISUAL 01	AUDITORY 02	HAPTIC 03	SMELL 04	TASTE 05	NON-SPECIFIC 06	SELF 07	FAMILY 08	PEERS 09	ADULTS 10	SCHOOL 11	COMMUNITY 12	ORAL LANGUAGE 13	WRITTEN LANGUAGE 14	2ND LANGUAGE 15	NUMBERS 16	WTS. & MEASURES 17	MONEY 18	TIME 19	BODY MOVEMENT 20	MUSIC 21	VISUAL ARTS 22	DRAMATIC ARTS 23	POETRY & LITERATURE 24	EARTH/UNIVERSE 25	PLANTS, ANIMALS 26	HEALTH, NUTRITION 27	BUSINESS & MANUF. 28	BUILDING 29	TRANSPORTATION 30	SOCIETY 31	UNSPECIFIED 99
REMEMBERING 101																																
DISCRIMINATING 102																				●												
SORTING 103																																
IDENTIFYING 104	●						●					●	●																			
RECOGNIZING SEQUENCES 105																																
UNDERSTANDING 106							●																									
PROBLEM SOLVING 107																																
CRITICAL THINKING 108																																●
CREATING, INVENTING 109							●													●		●	●	●								
UNSPECIFIED 199																																
ATTENDING 201							●																									
FEELING FREE, NOT FEARING 202							●	●	●	●			●	●						●	●	●	●			●	●					
ENJOYING, LIKING 203																				●		●										
EXPECTING SUCCESS 204																																
BELIEVING, VALUING 205																																
REFRAINING, RESPECTING 206							●		●	●			●																			
WILLING TO DO 207																																
BEING CURIOUS 208																																
INITIATING 209							●																									
EMPATHIZING 210								●	●	●																						
CHECKUPS 301																																
PHYSICAL STATUS 302																																
PHYSICAL SKILLS 303																																
ANALYZING 401																																
PLANNING 402																																
MANAGING 403																																
COPING/PROBLEMS 404																																
COPING/FEELINGS 405																																

income children will have a need of self-concept enhancement and (b) a prerequisite of high self-regard is positive identification with one's socially ascribed reference group".

Cross insists that, in fact, minority children in general have high overall self confidence upon entering Head Start, and that they retain this high global self-esteem during Head Start and subsequent schooling, at least during the age range of concern to this project. What appears to be independently nurtured is an academic specific self esteem, which is a function of Head Start and early school experiences. Cross therefore recommends:

> that in the proposed Head Start program-effects study, measures that tap only global self-esteem and general emotional functioning should be avoided. Measures that emphasize academic self-esteem and emotional functioning within academic contexts should be favored. The 'ideal' assessment techniques would incorporate self-esteem and emotional development scores for global and academic specific functioning...; if observations are made only in academic settings but are coded as 'general emotional development', we would run the risk of making our children appear to have a deficit in the emotional domain. A two-setting (familial and academic) approach would show the effects of Head Start as something which is added to the child's existing strengths (Cross, 1978, p. 20; emphases in original).

Accordingly, it is recommended that separate indices be developed for global and academic self-concept. A summary of this Dimension focuses on the global characteristics. The next discussed Dimension, "Attitudes towards Schooling" focuses on the academic.

The second common misconception that Cross (1978) points out is with regard to the inclusion of characteristics regarding ethnic identity with self-concept. After reviewing studies on the subject, Cross concludes that:

> the few studies that have actually administered separate measures of social and personal identity have, in general, failed to substantiate the long held belief that group identity is a major determining factor of self-esteem. If anything, it appears that the two domains are not related.... Under no circumstances should racial preference or self-identification techniques be used to assess self-concept, self-esteem or "general emotional development " (Cross, 1978, p. 20).

The knowledges and skills and attitudes and habits related to racial/ethnic background that were identifed in the input workshops as indicators of desired self concept are, nevertheless, included as recommended for measurement. Measures developers should be alerted to test Cross' hypotheses in the pilot testing activities of Phase II.

A complete listing of the characteristics included under "Self Concept" is provided on pages 30 through 36 of Appendix D. They are summarized by their process-content characteristics in Figure 14.

Attitudes Toward School. In accordance with Cross' recommendations on Self Concept, as discussed above, attitudes toward school were not included there; they are the focus of this Dimension. Included under "Attitudes Toward School" are the basic knowledges and skills prerequisite to being able to function in a school setting and,

Figure 14

Process and Content of Characteristics Identified in Dimension 63: SELF-CONCEPT
unvalidated

CONTENT

PROCESS	01 VISUAL	02 AUDITORY	03 HAPTIC	04 SMELL	05 TASTE	06 NON-SPECIFIC	07 SELF	08 FAMILY	09 PEERS	10 ADULTS	11 SCHOOL	12 COMMUNITY	13 ORAL LANGUAGE	14 WRITTEN LANGUAGE	15 2ND LANGUAGE	16 NUMBERS	17 WTS. & MEASURES	18 MONEY	19 TIME	20 BODY MOVEMENT	21 MUSIC	22 VISUAL ARTS	23 DRAMATIC ARTS	24 POETRY & LITERATURE	25 EARTH/UNIVERSE	26 PLANTS, ANIMALS	27 HEALTH, NUTRITION	28 BUSINESS & MANUF.	29 BUILDING	30 TRANSPORTATION	31 SOCIETY	99 UNSPECIFIED
REMEMBERING 101																																
DISCRIMINATING 102																																
SORTING 103																																
IDENTIFYING 104							●	●							●																	
RECOGNIZING SEQUENCES 105							●												●													
UNDERSTANDING 106							●	●															●									
PROBLEM SOLVING 107																																
CRITICAL THINKING 108							●																									
CREATING, INVENTING 109							●														●	●										
UNSPECIFIED 199																																
ATTENDING 201							●		●														●									
FEELING FREE, NOT FEARING 202						●	●	●	●	●			●	●	●							●	●									●
ENJOYING, LIKING 203							●	●												●	●	●		●								
EXPECTING SUCCESS 204							●			●			●		●																	
BELIEVING, VALUING 205							●	●																								
REFRAINING, RESPECTING 206							●	●																			●					
WILLING TO DO 207							●	●		●	●																					
BEING CURIOUS 208							●																									
INITIATING 209							●																									
EMPATHIZING 210																																
CHECKUPS 301																																
PHYSICAL STATUS 302																																
PHYSICAL SKILLS 303																																
ANALYZING 401																																
PLANNING 402																																
MANAGING 403																																
COPING/PROBLEMS 404																																
COPING/FEELINGS 405																																

therefore, to being able to feel positive about school experiences. Of particular emphasis are the habits that reflect positive attitudes to schooling and oneself in a school environment. (see process categories 206 through 210).

A complete listing of the characteristics identified as contributing to "Attitudes Toward School" is provided on pages 37 through 44 of Appendix D. The summary of characteristics by process-content properties is provided in Figure 15.

Independence. "Independence" has been selected as a separate Dimension to specifically focus on behaviors related to children being able to manage by themselves. This quality was particularly valued by Head Start parents and staff; it is also frequently identified as a separate construct by teachers and child development scholars. It was not identified in the input workshops as a specific area for discussion, however, resulting in relatively few characteristics being elicited.

Independence is defined as an attitude of self-reliance or of resistance to control by others (English & English, 1958). A low level of independence is associated with poor Self-Concept; an excessively high level with poor Relationships with Peers and Adults. Therefore this Dimension may be anticipated to interrelate with both these other Dimensions.

Independence may be an important Dimension to focus on separately in program-effects measurement because of its significant influence on these other Dimensions and also its amenability for targeting as a program objective. Whereas Self-Concept and Relationships cannot be directly taught, and are a second-order effect of intervention strategies, all of the characteristics of independence can be explicitly targeted and taught. As the characteristics indicative of independence often evidence, in addition, competence in a particular content area, many of the characteristics here identified are also indentified under the Dimensions addressing such other competencies. The importance of this Dimension should be reevaluated, therefore, during the pilot and field testing stages of this project.

The specific characteristics identified with "Independence" are listed on pages 45 and 46 of Appendix D; Figure 16 summarizes these characteristics by their process-content properties.

Sharing and Competing. Although the characteristics identified with "Sharing and Competing" could also be appropriately associated with the Dimension "Peer Relationships", next discussed, it is recommended that they be measured as a separate Dimension. Sharing and competing behaviors were the most frequently identified concerns of Head Start parents and staffs. For whatever reasons, it is important that this Dimension be measured as it is of great interest to those most intimately involved with Head Start children.

In fact, effective sharing and competing can have significant influence on many of the other Dimensions of development (e.g., Peer and Adult Relationships, Self Concept, Gross Motor Skills, Attitudes Toward Schooling) and on overall competence.

The characteristics describing development of Sharing and Competing are listed on pages 47 and 48 of Appendix D. They are also summarized by their process-content attributes in Figure 17.

Peer Relationships. Peer Relationships was considered by both input-workshop participants and scholars as an important Dimension for measurement. "Behavior

Figure 15

Process and Content of Characteristics Identified in Dimension 64: ATTITUDES TOWARD SCHOOL
unvalidated

CONTENT

PROCESS	VISUAL 01	AUDITORY 02	HAPTIC 03	SMELL 04	TASTE 05	NON-SPECIFIC 06	SELF 07	FAMILY 08	PEERS 09	ADULTS 10	SCHOOL 11	COMMUNITY 12	ORAL LANGUAGE 13	WRITTEN LANGUAGE 14	2ND LANGUAGE 15	NUMBERS 16	WTS. & MEASURES 17	MONEY 18	TIME 19	BODY MOVEMENT 20	MUSIC 21	VISUAL ARTS 22	DRAMATIC ARTS 23	POETRY & LITERATURE 24	EARTH/UNIVERSE 25	PLANTS, ANIMALS 26	HEALTH, NUTRITION 27	BUSINESS & MANUF. 28	BUILDING 29	TRANSPORTATION 30	SOCIETY 31	UNSPECIFIED 99
REMEMBERING 101											●																					
DISCRIMINATING 102																																
SORTING 103																																
IDENTIFYING 104							●	●			●					●																
RECOGNIZING SEQUENCES 105											●		●																			●
UNDERSTANDING 106	●										●																					
PROBLEM SOLVING 107											●																					●
CRITICAL THINKING 108							●																									●
CREATING, INVENTING 109																																●
UNSPECIFIED 199		●									●																					
ATTENDING 201										●	●		●	●							●	●	●	●	●	●						
FEELING FREE, NOT FEARING 202							●	●	●	●	●			●	●	●										●						
ENJOYING, LIKING 203							●				●			●							●	●	●			●						
EXPECTING SUCCESS 204														●																		
BELIEVING, VALUING 205												●									●	●				●						
REFRAINING, RESPECTING 206									●		●		●		●											●						
WILLING TO DO 207											●			●																		●
BEING CURIOUS 208											●			●		●									●	●		●				●
INITIATING 209							●				●			●																		
EMPATHIZING 210											●															●						
CHECKUPS 301																																
PHYSICAL STATUS 302																																
PHYSICAL SKILLS 303																																
ANALYZING 401																																
PLANNING 402																																
MANAGING 403																																
COPING/PROBLEMS 404																																
COPING/FEELINGS 405																																

Figure 16

Process and Content of Characteristics Identified in Dimension 65: INDEPENDENCE unvalidated

CONTENT

PROCESS	VISUAL 01	AUDITORY 02	HAPTIC 03	SMELL 04	TASTE 05	NON-SPECIFIC 06	SELF 07	FAMILY 08	PEERS 09	ADULTS 10	SCHOOL 11	COMMUNITY 12	ORAL LANGUAGE 13	WRITTEN LANGUAGE 14	2ND LANGUAGE 15	NUMBERS 16	WTS. & MEASURES 17	MONEY 18	TIME 19	BODY MOVEMENT 20	MUSIC 21	VISUAL ARTS 22	DRAMATIC ARTS 23	POETRY & LITERATURE 24	EARTH/UNIVERSE 25	PLANTS, ANIMALS 26	HEALTH, NUTRITION 27	BUSINESS & MANUF. 28	BUILDING 29	TRANSPORTATION 30	SOCIETY 31	UNSPECIFIED 99
REMEMBERING 101																																
DISCRIMINATING 102																																
SORTING 103																																
IDENTIFYING 104																																
RECOGNIZING SEQUENCES 105																																
UNDERSTANDING 106							●																									
PROBLEM SOLVING 107																																
CRITICAL THINKING 108							●																									
CREATING, INVENTING 109																																
UNSPECIFIED 199																																
ATTENDING 201																																
FEELING FREE, NOT FEARING 202							●	●		●	●																●					
ENJOYING, LIKING 203								●																			●					
EXPECTING SUCCESS 204																																
BELIEVING, VALUING 205																																
REFRAINING, RESPECTING 206								●																			●					
WILLING TO DO 207							●																									
BEING CURIOUS 208																																
INITIATING 209							●																									
EMPATHIZING 210																																
CHECKUPS 301																																
PHYSICAL STATUS 302																																
PHYSICAL SKILLS 303																																
ANALYZING 401																																
PLANNING 402																																
MANAGING 403																																
COPING/PROBLEMS 404																																
COPING/FEELINGS 405																																

Figure 17

Process and Content of Characteristics
Identified in Dimension 66: SHARING & COMPETING
unvalidated

CONTENT

PROCESS	01 VISUAL	02 AUDITORY	03 HAPTIC	04 SMELL	05 TASTE	06 NON-SPECIFIC	07 SELF	08 FAMILY	09 PEERS	10 ADULTS	11 SCHOOL	12 COMMUNITY	13 ORAL LANGUAGE	14 WRITTEN LANGUAGE	15 2ND LANGUAGE	16 NUMBERS	17 WTS. & MEASURES	18 MONEY	19 TIME	20 BODY MOVEMENT	21 MUSIC	22 VISUAL ARTS	23 DRAMATIC ARTS	24 POETRY & LITERATURE	25 EARTH/UNIVERSE	26 PLANTS, ANIMALS	27 HEALTH, NUTRITION	28 BUSINESS & MANUF.	29 BUILDING	30 TRANSPORTATION	31 SOCIETY	99 UNSPECIFIED
REMEMBERING 101																																
DISCRIMINATING 102																																
SORTING 103																																
IDENTIFYING 104													●																			
RECOGNIZING SEQUENCES 105																																
UNDERSTANDING 106									●																							
PROBLEM SOLVING 107																																
CRITICAL THINKING 108																																
CREATING, INVENTING 109																																
UNSPECIFIED 199																																
ATTENDING 201									●																							
FEELING FREE, NOT FEARING 202							●		●																							
ENJOYING, LIKING 203																				●												
EXPECTING SUCCESS 204																																
BELIEVING, VALUING 205							●																									
REFRAINING, RESPECTING 206								●	●																							
WILLING TO DO 207									●																							
BEING CURIOUS 208																																
INITIATING 209									●																							
EMPATHIZING 210								●	●	●	●																					
CHECKUPS 301																																
PHYSICAL STATUS 302																																
PHYSICAL SKILLS 303																																
ANALYZING 401																																
PLANNING 402																																
MANAGING 403																																
COPING/PROBLEMS 404																																
COPING/FEELINGS 405																																

towards peers is universally accepted as an important aspect of child development" (Katz & Jacobson, 1978).

The characteristics included in Peer Relationships encompass verbal and non-verbal communications, being interested in establishing and maintaining relationships with others, sharing and cooperating, being able to resolve problems and quarrels, expressing empathy, and initiating actions to enhance friendships. As such, they draw upon and overlap competencies addressed by other Dimensions.

The specific characteristics are listed on pages 49 - 54 of Appendix D; they are also summarized in Figure 18.

Adult Relationships. Adult Relationships involve characteristics evidencing both children's ability to communicate and get along with adults and children's understanding and identity of adult roles. As such, it also represents a large area of desired development and a most important Dimension for measurement.

Adult Relationships also has one of the largest number of characteristics identifed as indicators of desired development. The specific characteristics are listed on pages 55 - 63 of Appendix D; they are also summarized by their process -content properties in Figure 19.

Health and Physical Development

It is generally recognized that the health and physical development of young children is fundamental to their effective functioning in all other domains of behavior. This is evident in the major emphasis given to "Health" and "Nutrition" in Head Start Program Performance Standards (1975b, esp. pp. 16-54). It is also evident in the hundreds of health-related objectives identified for this project by scholars and by parents and teachers who participated in the input workshops. Thus, a critical basis for evaluating Head Start and other early childhood programs is their measured impact on the health and physical development of young children.

The Domain of Health and Physical Development, as conceived in this project, encompasses those physiological developments and learnings primarily associated with soundness of body. It is defined by an overall state of "wellness", which includes such characteristics as optimal functioning of bodily organs, resistance to and resilience from sickness or injury, stamina, physical strength and dexterity.

As panelist Eichorn points out in a paper prepared for this project (Eichorn, 1978), being healthy involves the attainment of specific levels of development beyond which "may either not be particularly helpful or, indeed, occasionally dangerous". Heart rate, blood pressure, visual and auditory acuity, and muscular development are examples of characteristics that are optimally defined for children within specific limits. There are similar limits to the desirable effects of learnings within the Domain, e.g., the concern of children for their own health and safety. Thus, this Domain, together with each of its Dimensions and characteristics, is conceptualized as involving development in an optimum range of conditions within high and low limits. This is in contrast to the Cognitive Domain in which there are no conceptual limits to desired development of children.

Figure 18

Process and Content of Characteristics Identified in Dimension 67: PEER RELATIONSHIPS unvalidated

CONTENT

PROCESS	VISUAL 01	AUDITORY 02	HAPTIC 03	SMELL 04	TASTE 05	NON-SPECIFIC 06	SELF 07	FAMILY 08	PEERS 09	ADULTS 10	SCHOOL 11	COMMUNITY 12	ORAL LANGUAGE 13	WRITTEN LANGUAGE 14	2ND LANGUAGE 15	NUMBERS 16	WTS. & MEASURES 17	MONEY 18	TIME 19	BODY MOVEMENT 20	MUSIC 21	VISUAL ARTS 22	DRAMATIC ARTS 23	POETRY & LITERATURE 24	EARTH/UNIVERSE 25	PLANTS, ANIMALS 26	HEALTH, NUTRITION 27	BUSINESS & MANUF. 28	BUILDING 29	TRANSPORTATION 30	SOCIETY 31	UNSPECIFIED 99
REMEMBERING 101																																
DISCRIMINATING 102									●																							
SORTING 103																																
IDENTIFYING 104									●																							
RECOGNIZING SEQUENCES 105									●																							
UNDERSTANDING 106									●																							
PROBLEM SOLVING 107									●																							
CRITICAL THINKING 108																																
CREATING, INVENTING 109																																
UNSPECIFIED 199																																
ATTENDING 201									●																							
FEELING FREE, NOT FEARING 202									●						●					●												
ENJOYING, LIKING 203									●		●																					
EXPECTING SUCCESS 204																																
BELIEVING, VALUING 205									●																							
REFRAINING, RESPECTING 206									●				●																			
WILLING TO DO 207									●																							
BEING CURIOUS 208																																
INITIATING 209									●																							
EMPATHIZING 210									●																							
CHECKUPS 301																																
PHYSICAL STATUS 302																																
PHYSICAL SKILLS 303																																
ANALYZING 401																																
PLANNING 402																																
MANAGING 403																																
COPING/PROBLEMS 404																																
COPING/FEELINGS 405																																

Figure 19

Process and Content of Characteristics Identified in Dimension 68: ADULT RELATIONSHIPS unvalidated

CONTENT

PROCESS	01 VISUAL	02 AUDITORY	03 HAPTIC	04 SMELL	05 TASTE	06 NON-SPECIFIC	07 SELF	08 FAMILY	09 PEERS	10 ADULTS	11 SCHOOL	12 COMMUNITY	13 ORAL LANGUAGE	14 WRITTEN LANGUAGE	15 2ND LANGUAGE	16 NUMBERS	17 WTS. & MEASURES	18 MONEY	19 TIME	20 BODY MOVEMENT	21 MUSIC	22 VISUAL ARTS	23 DRAMATIC ARTS	24 POETRY & LITERATURE	25 EARTH/UNIVERSE	26 PLANTS, ANIMALS	27 HEALTH, NUTRITION	28 BUSINESS & MANUF.	29 BUILDING	30 TRANSPORTATION	31 SOCIETY	99 UNSPECIFIED
REMEMBERING 101																																
DISCRIMINATING 102								●																								
SORTING 103																																
IDENTIFYING 104								●		●		●																				
RECOGNIZING SEQUENCES 105																																
UNDERSTANDING 106								●		●		●	●																		●	
PROBLEM SOLVING 107																																
CRITICAL THINKING 108										●																						
CREATING, INVENTING 109																																
UNSPECIFIED 199																																
ATTENDING 201								●		●		●																				
FEELING FREE, NOT FEARING 202								●		●	●	●	●																			
ENJOYING, LIKING 203								●		●		●																				
EXPECTING SUCCESS 204										●																						
BELIEVING, VALUING 205							●	●				●																			●	
REFRAINING, RESPECTING 206							●	●		●	●	●	●																		●	
WILLING TO DO 207							●	●		●																						
BEING CURIOUS 208										●		●																				
INITIATING 209										●																						
EMPATHIZING 210								●	●	●		●																				
CHECKUPS 301																																
PHYSICAL STATUS 302																																
PHYSICAL SKILLS 303																																
ANALYZING 401																																
PLANNING 402																																
MANAGING 403																																
COPING/PROBLEMS 404																																
COPING/FEELINGS 405																																

Within the Domain of Health and Physical Development, five interrelated Dimensions have been defined Health and Safety, Dental, Nutritional, Gross Motor and Fine Motor Development corresponding to areas of discrete program emphases. This classification is based on grounds of custom and convenience for practitioners, curriculum planners and developers, and others studying program effects on children. These Dimensions are very much interdependent, as is evidenced in the discussion on each of their characteristics.

Mental health is often included within a domain of health and physical development. In this project, mental health concerns are dealt with to some extent in the Domain of Social-Emotional Development, which is closely related to physical health. Low levels of social-emotional development may be appropriately described as unhealthy or sick conditions, resembling conditions of physical sickness in symptoms and treatments. Similarly, desirable levels of social-emotional development may be important contributors to physical health. Such relationships may be examined empirically through data from the proposed measures.

The determinants of attainment in each of the Dimensions of Health and Physical Development are hypothesized as including cognitive (knowledge) and affective (attitudes) characteristics, as well as the more commonly identified physiological attributes. This is in recognition of the importance of knowledge, and particularly appropriate habits, to higher levels of development within the Domain. Program effects on health-related knowledge and habits have not been included in previous evaluations of Head Start, even though they are an important curriculum component of the program. Such health-related cognitive and affective learnings are hypothesized here as significant determinants of children's development in the several Dimensions of the Domain and to the attainment of social competence. The characteristics comprising the Domain are summarized in Figure 20.

A somewhat unique problem is encountered in recommending characteristics for measurement in the Domain of Health and Physical Development. In the Domains discussed previously, the major concern is selecting and defining constructs, i.e., the Dimensions to be measured, characteristics defining the respective Dimensions become the characteristics recommended for their measurement. In the area of Health and Physical Development, however, the constructs, once selected, are relatively readily understood and defined. In fact, there are many more characteristics associated with each Dimension than are desired for measures of program effects. Most of the physiological characteristics, however, develop to a significant degree without any program influence; for example, children develop physically, and develop many indications of sound nutrition regardless of a program's intervention. The challenge in Health and Physical Development is, therefore, to recommend characteristics that are particularly affected by the program, i.e., to focus on the areas on which Head Start concentrates and makes the difference. Characteristics of the knowledges-and-skills and attitudes-and habits types are hypothesized as providing this special sensitivity. The discussion of the characteristics to be measured under each of the Dimensions describes other considerations on which the recommendations to measures developers are based.

Health and Safety. The Dimension of Health and Safety is conceptualized as consisting of three types of child-development characteristics: (1) physiological conditions and services that relate to wellness, e.g., the development of appropriate height and weight and having regular medical check-ups; (2) attitudes and habits related to health and safety; and (3) knowledges and skills contributing to wellness and which, presumably, affect physiological status. The boundaries of the medical

Figure 20

Process and Content of Characteristics Identified in HEALTH & PHYSICAL DEVELOPMENT unvalidated

CONTENT

PROCESS	VISUAL 01	AUDITORY 02	HAPTIC 03	SMELL 04	TASTE 05	NON-SPECIFIC 06	SELF 07	FAMILY 08	PEERS 09	ADULTS 10	SCHOOL 11	COMMUNITY 12	ORAL LANGUAGE 13	WRITTEN LANGUAGE 14	2ND LANGUAGE 15	NUMBERS 16	WTS. & MEASURES 17	MONEY 18	TIME 19	BODY MOVEMENT 20	MUSIC 21	VISUAL ARTS 22	DRAMATIC ARTS 23	POETRY & LITERATURE 24	EARTH/UNIVERSE 25	PLANTS, ANIMALS 26	HEALTH, NUTRITION 27	BUSINESS & MANUF. 28	BUILDING 29	TRANSPORTATION 30	SOCIETY 31	UNSPECIFIED 99
REMEMBERING 101																																
DISCRIMINATING 102																																
SORTING 103																																
IDENTIFYING 104								●		●				●		●										●	●				●	
RECOGNIZING SEQUENCES 105																										●	●					
UNDERSTANDING 106										●																	●					
PROBLEM SOLVING 107																																
CRITICAL THINKING 108																																
CREATING, INVENTING 109																																
UNSPECIFIED 199																																
ATTENDING 201																										●	●					
FEELING FREE, NOT FEARING 202							●			●										●							●					
ENJOYING, LIKING 203																				●							●					
EXPECTING SUCCESS 204							●													●												
BELIEVING, VALUING 205																																
REFRAINING, RESPECTING 206								●		●	●	●								●						●	●				●	
WILLING TO DO 207								●												●							●					
BEING CURIOUS 208							●					●																				
INITIATING 209																											●					
EMPATHIZING 210																																
CHECKUPS 301																											●					
PHYSICAL STATUS 302	●	●				●	●																				●					
PHYSICAL SKILLS 303			●											●						●	●	●										
ANALYZING 401																																
PLANNING 402																																
MANAGING 403																																
COPING/PROBLEMS 404																																
COPING/FEELINGS 405																																

Dimension might be viewed as including development also in the Dental and Nutritional Dimensions. However, since each of them relates to a specialized aspect of health with particular objectives, they are considered here as separate Dimensions.

The importance of measuring development toward "wellness" in the Dimension of Health and Safety, as opposed to measuring only the absence of sickness, is recognized by previous studies in this area (e.g., American Associations of Psychiatric Services, 1978; Raizen and Bobrow, 1974, Stebbins, et al., 1978); but their final recommendations are for gathering data thorugh standard medical examinations and developmental assessments. Although acknowledging important differences between absence of sickness and wellness, they chose to assess the former because of limitations in available measuring technique. This project seeks to overcome such limitations.

Regular medical examinations, innoculations and appropriate treatments are technically services rather than child characteristics. They are, however, conceptualized as represnting important resources attained by children analagous to skills in the cognitive domain, and are, therefore, included as identified in the input-workshops. They represent assurances of the present and continued satisfactory status of specific physiological attributes. The specific attributes recommended (Gordon, 1978) to be included under these services are listed in Table VII.

The other characteristics identified as important to the Dimension of Health and Safety are listed in pages 64 to 67 of Appendix D. They are also summarized in Figure 21 by their process-content properties.

A particular area of development omitted from the list of characteristics appears worthy of special notation; it has to do with children's ability to function under a wide range of stress-producing environments. It is hypothesized that such a "functional stress range" could contribute to increased competency in all areas, for example, in the ability to concentrate on learning tasks, in the willingness to undertake risks and responsibilities, and in the ability to resist disease and prolonged physical strain. It has not been recommended since clinical and experimental measuring techniques are presently considered unsuitable for use in program evaluation. It is recommended that the area receive ongoing attention and consideration for inclusion when the state-of-the-art conditions appear to be appropriately developed. These issues are more completely explicated in Perspectives on the Measurement of Functional Stress Range (Taub, 1980).

Dental. The structure of the Dental Dimension is conceived much like that of the Medical Dimension. All three types of characteristics are identified as indices of program effect: (1) physiological dental statuses, (2) dental related attitudes and habits, and (3) knowledges and skills. The characteristics in the Dental Dimension that were identified as desirable by participants in the Head Start Input Workshops are listed in pages 68 and 69 of Appendix D and are summarized in Figure 22. Included as most important among these characteristics are regular dental exmainations and appropriate treatments. The attributes recommended (Gordon, 1978) for attention through these services are as follows:

TABLE VII

Attributes Recommended For Check-Up and Treatment Through Medical Examination In The Health And Safety Dimension

	Attributes To Be Checked:	Measuring Techniques:
1.	Blood pressure	Sphygmomanometer, etc.
2.	Respiratory and heart disorders	Aysculation (e.g., through use of stethoscope)
3.	Speech articulation (neurologically or anatomically based)	Denver Articulation Screening Test
4.	Visual acuity (abnormality or excessive refractory error)	3-4 year olds: Allen Osterburg Picture Codes 5-7 year olds: Snellen Illiterate E Test, or Titmus test
5.	Auditory acuity	Audiometric and Impedence Testing
6.	Blood disorders--iron deficiency; glucose--6-phosphate-deghydrosunose (G6PD) deficiency; sickle cell disease (for children) with African or Mediterranean background	Standard blood test, coordinated with requirements in nutritional dimension
7.	Urinary tract infection (bacteruria), for girls only	Dip-slide urine culture
8.	Tuberculosis	Tine Test
9.	Other physical findings	Physical examination
10.	Innoculations: a. Diptheria-Tetanus and Pertussis Vaccine (DTP) -- 4 doses required: b. Trivalent Polio vaccine-- 3 doses required; and c. Measles, mumps and rubella vaccine (MMR), separately or combined	Inspection of medical records

Figure 21

Process and Content of Characteristics Identified in Dimension 71: HEALTH AND SAFETY unvalidated

CONTENT

PROCESS	VISUAL 01	AUDITORY 02	HAPTIC 03	SMELL 04	TASTE 05	NON-SPECIFIC 06	SELF 07	FAMILY 08	PEERS 09	ADULTS 10	SCHOOL 11	COMMUNITY 12	ORAL LANGUAGE 13	WRITTEN LANGUAGE 14	2ND LANGUAGE 15	NUMBERS 16	WTS. & MEASURES 17	MONEY 18	TIME 19	BODY MOVEMENT 20	MUSIC 21	VISUAL ARTS 22	DRAMATIC ARTS 23	POETRY & LITERATURE 24	EARTH/UNIVERSE 25	PLANTS, ANIMALS 26	HEALTH, NUTRITION 27	BUSINESS & MANUF. 28	BUILDING 29	TRANSPORTATION 30	SOCIETY 31	UNSPECIFIED 99
REMEMBERING 101																																
DISCRIMINATING 102																																
SORTING 103																																
IDENTIFYING 104																										●	●				●	
RECOGNIZING SEQUENCES 105																											●					
UNDERSTANDING 106										●																	●					
PROBLEM SOLVING 107																																
CRITICAL THINKING 108																																
CREATING, INVENTING 109																																
UNSPECIFIED 199																																
ATTENDING 201																										●						
FEELING FREE, NOT FEARING 202																											●					
ENJOYING, LIKING 203																				●							●					
EXPECTING SUCCESS 204																																
BELIEVING, VALUING 205																																
REFRAINING, RESPECTING 206										●										●						●	●				●	
WILLING TO DO 207																											●					
BEING CURIOUS 208							●					●																				
INITIATING 209																											●					
EMPATHIZING 210																																
CHECKUPS 301																											●					
PHYSICAL STATUS 302	●	●				●	●																				●					
PHYSICAL SKILLS 303																																
ANALYZING 401																																
PLANNING 402																																
MANAGING 403																																
COPING/PROBLEMS 404																																
COPING/FEELINGS 405																																

Figure 22

Process and Content of Characteristics Identified in Dimension 72: DENTAL unvalidated

CONTENT

PROCESS

	VISUAL 01	AUDITORY 02	HAPTIC 03	SMELL 04	TASTE 05	NON-SPECIFIC 06	SELF 07	FAMILY 08	PEERS 09	ADULTS 10	SCHOOL 11	COMMUNITY 12	ORAL LANGUAGE 13	WRITTEN LANGUAGE 14	2ND LANGUAGE 15	NUMBERS 16	WTS. & MEASURES 17	MONEY 18	TIME 19	BODY MOVEMENT 20	MUSIC 21	VISUAL ARTS 22	DRAMATIC ARTS 23	POETRY & LITERATURE 24	EARTH/UNIVERSE 25	PLANTS, ANIMALS 26	HEALTH, NUTRITION 27	BUSINESS & MANUF. 28	BUILDING 29	TRANSPORTATION 30	SOCIETY 31	UNSPECIFIED 99
REMEMBERING 101																																
DISCRIMINATING 102																																
SORTING 103																																
IDENTIFYING 104										●																	●					
RECOGNIZING SEQUENCES 105																											●					
UNDERSTANDING 106																											●					
PROBLEM SOLVING 107																																
CRITICAL THINKING 108																																
CREATING, INVENTING 109																																
UNSPECIFIED 199																																
ATTENDING 201																											●					
FEELING FREE, NOT FEARING 202										●																	●					
ENJOYING, LIKING 203																											●					
EXPECTING SUCCESS 204																																
BELIEVING, VALUING 205																																
REFRAINING, RESPECTING 206																											●					
WILLING TO DO 207																											●					
BEING CURIOUS 208																																
INITIATING 209																																
EMPATHIZING 210																																
CHECKUPS 301																											●					
PHYSICAL STATUS 302																																
PHYSICAL SKILLS 303			●																													
ANALYZING 401																																
PLANNING 402																																
MANAGING 403																																
COPING/PROBLEMS 404																																
COPING/FEELINGS 405																																

Attribute	Measuring Techniques
1. Decayed-extracted-filled surface status	DEF's examination
2. Periodontal health status	P-M-A index of gingivitis
3. Plaque status	Podshadley Oral Hygiene Index (PHOI)

This list requires further comment and recommendation to guide measures developers.

The incidence of tooth decay in children 3 to 7 years old has been found consistently to be significantly greater among low-income and non-white children than among other children. If neglected, decay can result in severe pain, permanent loss of teeth, loss of mouth functioning, and a deleterious effect on overall health. The identification and treatment of decay and prevention of the loss of teeth is required by the *Head Start Program Performance Standards*. A decayed-extracted-filled (DEF) examination, utilizing standard techniques for inspecting dental status, can be used to measure program effects on this important factor.

Although the *Program Performance Standards* require that all children receive dental examinations, dental records cannot be assumed to report caries on a common basis. Early lesions, in particular, can easily be ignored; or what some consider enamel defects unrelated to dental caries can be reported as clincial caries by others. A system of graded measures is required which can be used reliably by data gatherers. Such a system can be sensitive to program-effect changes by identifying caries that are in the early stages (e.g., white-spot stage), and measuring their greater or lesser development over the period of the program. Such systems have been developed for use in longitudinal studies (Edgar, Jenkins, and Rugg-Gunn, Unpublished; Fehr, 1961; Marthaler, 1966), and are recommended for adoption and use as part of the battery of Head Start measures.

The mere presence or absence of defective surfaces is not in itself an adequate measure of program effects, since requirements for dental services normally increase over the time of the program at rates that vary with regional, racial/ethnic and other facors apparently independent of program content. A valid measure of program effects has to take into account both the potential need for services and the extent of services provided. Measurement procedures must also differentiate between loss of teeth as a result of corrective extractions and losses due to other causes. These factors and their relative importance for the index of Dental health are to be determined empirically during the pilot phase of the project.

Unattended peridontal disease can utlimately lead to loss of teeth and jaw bone, and to severely restricted functioning of the mouth and reduced overall effectiveness. Thus, reduction in the extent and severity of gum tissue infection is also an appropriate factor in the measurement of program effects in the dental dimension.

The measurement of caries-potential calls for biochemical techniques for which this project is not yet ready to offer a recommendation. A number of new techniques for this purpose are currently being researched (Bibby, 1975; Edgar, 1975; Eilen, 1976). The most promising of these techniques involves measuring the acid production in plaque (reduction in pH) after rinsing the mouth with a standard concentration carbohydrate solution. At this writing, no technique has been identified that is sensitive to a change brought about by the program; but there may be new developments that will serve this purpose.

Malocclusion in deciduous teeth of children between the ages of 3 and 7 may predict severity of conditions in permanent teeth, and thus may be useful in evaluations of individual children and in counseling parents on the need for future care. However, as no program treatment for children is specified for this condition, it is not included among the characteristics recommended to be measured.

The existence of plaque is an indication of the quality of self-administered oral hygiene, which is a targeted program effect. Plaque is also a contributor to periodontal disease. However, plaque development varies greatly with individuals as a function of factors that are not affected by programs; and therefore program-effects measures have to consider the potential for plaque without program intervention. Because of plaque's potential importance as a contributor to Dental health, measurement of the extent of its existence is recommended. However, its use as a separate characteristic and its weight in the overall index of Dental health, if any, needs to be evaluated during the pilot phase of the project.

Any effect on a child's potential for future development of caries is an important influence on Dental health. Head Start Program Performance Standards recognize this factor in recommending the application of topical fluoride in communities which lack adequate fluoride levels in the public water supply. The extent of treatment of teeth with fluorides, in the case of those children who live in communities without fluoridated water, should be documented as a program effect, the same as with innoculations in the Medical Dimension.

The physiological data-gathering techniques recommended for development in the Dental Dimension should be reviewed with the National Institute of Dental Research during the first year of the pilot phase of this project. New developments in the state-of-the-art may warrant revised recommendations.

Nutrition. Nutrition is a most important influence on children's overall health and physical development. The Head Start Program Performance Standards (OCD, 1975b; p. 39), therefore, requires the "provision of a high quality nutrition component". This component should provide food which will help meet the child's daily nutritional needs, should help children (and their families) understand the relationship of nutrition to overall health and competence) and should help children (and their families) develop the knowledge, skills, attitudes and habits to maintain a sound nutritional status, during and after leaving the Head Start Program. All of these desired outcomes need to be addressed by measures of program effects.

Nutrition also influences all the other Dimensions in the Domain -- Health, Dental, and Fine and Gross Motor Development; as such, nutrition is a vital part of overall preventative health care and wellness. Although there is a time lag between change in nutritional practices and a measurable change in physiological indices, the long term indicators of poor development in these other dimensions -- incidence of sickness, loss of teeth, etc. -- are also indicators of overall Nutritional development.

This confounding of effects presents a problem for measurement. In an article entitled "The Perilous Journey of Nutrition Evaluation", James E. Austin writes:

> Measuring changes in nutritional status via biochemical, clinical anthropometric, or dietary procedures is technically feasible. However, the degree of precision

and accuracy of the measuring instruments make one less than entirely comfortable. Important changes in the population's well-being may occur, but our evaluative tools may not be able to pick those out. Further complicating the task is the need for careful design of the evaluation's sample and data collection process that require expertise that is often not readily available. Furthermore, the meaning of the results is not fully understood because of a lack of knowledge about underlying functional relationships. For example, it is not clear what the significance is of receiving 80 versus 90% of one's recommended daily calorie allowance. Thus, the nature of malnutrition, the imperfections in the measuring instruments, the sophistication requirements of evaluation design, and the ambiguity of the data collected, have all combined to raise a technical barrier to evaluation (Austin, 1978).

Over the long term, the data available through the measures developed in this project may provide insights into some of the problems identified by Austin. For the short term, we may proceed with measures of the more sensitive characteristics hypothesized as associated with nutritional status. The aggregaton of data on Nutrition into a higher-level index of, say, Health and Physical Development, however, may be difficult or impossible to accomplish with statistical validity using currently available knowledge and techniques.

The characteristics identified as determinants of nutritional status by the participants in the input-workshops are listed on pages 70 - 72 of Appendix D. They are also summarized in Figure 23. They particularly address the knowledges and skills and attitudes and habits that are affected by Head Start and that should be more sensitive indicators of program effects than the physiological attitudes commonly measured.

Many of the recommended characteristics involve food preferences and habits that are particularly sensitive to cultural and regional influences. It is essential therefore, that the measures to be developed allow for wide diversity in the specific practices being assessed. The values implicit or explicit in such measures should not communicate negative judgments on any existing practices or preferences. Measuring techniques should accommodate all practices and preferences that are nutritionally sound and that reflect the diversity in eating customs among Head Start families.

Gross Motor Development. The importance of motor skills for effective functioning is generally recognized among scholars and practitioners in early childhood development, and this is notably true in the case of Head Start. As panelist Dorothy Eichorn states in a paper prepared for this project:

> Motor skills and coordination are an important component of competent performance and healthy development throughout the life span. Nevertheless, the younger the child the more prominent the role of the motor component in overall development. Although the motor component in self-help skills (e.g., dressing; toothbrushing) and in basic academic skills (e.g., writing, eye-hand coordination in transfer from blackboard to desk) is widely recognized, it is less widely recognized that speech involves the most complex motor skills evidenced by human beings. Because motor skills and coordination are basic to competence in a broad range of performances essential to everyday living as well as academic accomplishments, and because their exercise also promotes physical health, enhancement of motor development has always been one of the goals of Head Start programs (Eichorn, 1978, p. 1).

Figure 23

Process and Content of Characteristics Identified in Dimension 73: NUTRITIONAL unvalidated

CONTENT

PROCESS	01 VISUAL	02 AUDITORY	03 HAPTIC	04 SMELL	05 TASTE	06 NON-SPECIFIC	07 SELF	08 FAMILY	09 PEERS	10 ADULTS	11 SCHOOL	12 COMMUNITY	13 ORAL LANGUAGE	14 WRITTEN LANGUAGE	15 2ND LANGUAGE	16 NUMBERS	17 WTS. & MEASURES	18 MONEY	19 TIME	20 BODY MOVEMENT	21 MUSIC	22 VISUAL ARTS	23 DRAMATIC ARTS	24 POETRY & LITERATURE	25 EARTH/UNIVERSE	26 PLANTS, ANIMALS	27 HEALTH, NUTRITION	28 BUSINESS & MANUF.	29 BUILDING	30 TRANSPORTATION	31 SOCIETY	99 UNSPECIFIED
REMEMBERING 101																																
DISCRIMINATING 102																																
SORTING 103																																
IDENTIFYING 104								●																			●					
RECOGNIZING SEQUENCES 105																										●	●					
UNDERSTANDING 106																											●					
PROBLEM SOLVING 107																																
CRITICAL THINKING 108																																
CREATING, INVENTING 109																																
UNSPECIFIED 199																																
ATTENDING 201																																
FEELING FREE, NOT FEARING 202																											●					
ENJOYING, LIKING 203																											●					
EXPECTING SUCCESS 204																																
BELIEVING, VALUING 205																																
REFRAINING, RESPECTING 206								●																			●					
WILLING TO DO 207								●																			●					
BEING CURIOUS 208																																
INITIATING 209																											●					
EMPATHIZING 210																																
CHECKUPS 301																																
PHYSICAL STATUS 302																											●					
PHYSICAL SKILLS 303																																
ANALYZING 401																																
PLANNING 402																																
MANAGING 403																																
COPING/PROBLEMS 404																																
COPING/FEELINGS 405																																

The Head Start Program Performance Standards (p. 8) require the "promotion of physical growth by activities designed to develop: large muscles (wheel toys, climbing apparatus, blocks); small muscles (scissors, clay, puzzles, small blocks); eye-hand coordinaton (puzzles, balls, lotto); body awareness; rhythm and movement (dancing, musical instruments)". This requirement emphasizes the importance of motor skills development in the assessment of Head Start Program effects.

Most motor skills of concern to early-childhood programs represent the coordination of a perception, usually visual or auditory, with a bodily action. These are often referred to as perceptual-motor skills, which represent essentially the ability to have one's muscles do what one wants and intends to do. "These are important for many reasons. First, the children's muscular development is essential for exercise and physical health. A child who moves awkwardly is likely to develop feelings of inferiority, and if he is preoccupied with his awkwardness, he is not free to concentrate on more intellectual tasks. There is also a body of evidence which suggests that peceptual motor disabilities may be associated with various learning problems. Perceptual-motor objectives are usually divided into those involving gross muscle coordination and those involving fine motor coordination" (Kamii, 1971). The customary distinction between gross motor and fine motor skills is maintained in these recommendations. Their development in a child occurs at different stages and involves different muscles, nerve centers and teaching strategies.

As children develop gross motor skills to a relatively high degree, independent of a program intervention, it is recommended that those characteristics involving the more complex motor skills and higher degrees of coordination, integration, control and efficiency should be focused on by measures developers (Eichorn, 1978). Also the knowledges and skills and attitudes and habits which are not explicitly learned in extra-program experiences are particularly recommended for measurement.

The characteristics recommended in the input-workshops as indications of Gross Motor Development are listed on pages 73 - 76 of Appendix D; they are summarized in Figure 24. These recommendations are designed to avoid the problems of measures used in past evaluations of motor development among Head Start children which have found, in general, no significant differences between program participants and control groups (Eichorn, 1978).

Fine Motor Skills. The Dimension of Fine Motor Development is conceptualized to be very similar in structure to Gross Motor Development. In both cases, development is evidenced by changes in knowledge, attitudes and habits, and physical skills. The specific charcteristics identified in the input-workshops as associated with Fine Motor Skills are listed on pages 77 - 79 of Appendix D; they are also summarized in Figure 25.

In addition to these characteristics, the perceptual-motor skills required in speech articulation, identified elsewhere by the participants, are proposed to be measured as part of the Fine Motor Dimension. Their development involves internal processes, learning experiences, and teaching strategies more closely associated with motor skills than with any others. These perceptual-motor skills are prerequisite to the more advanced development of Langauge Skills proposed as a Dimension for measurement in the Cognitive Domain.

Figure 24

Process and Content of Characteristics Identified in Dimension 74: GROSS MOTOR unvalidated

CONTENT

PROCESS	VISUAL 01	AUDITORY 02	HAPTIC 03	SMELL 04	TASTE 05	NON-SPECIFIC 06	SELF 07	FAMILY 08	PEERS 09	ADULTS 10	SCHOOL 11	COMMUNITY 12	ORAL LANGUAGE 13	WRITTEN LANGUAGE 14	2ND LANGUAGE 15	NUMBERS 16	WTS. & MEASURES 17	MONEY 18	TIME 19	BODY MOVEMENT 20	MUSIC 21	VISUAL ARTS 22	DRAMATIC ARTS 23	POETRY & LITERATURE 24	EARTH/UNIVERSE 25	PLANTS, ANIMALS 26	HEALTH, NUTRITION 27	BUSINESS & MANUF. 28	BUILDING 29	TRANSPORTATION 30	SOCIETY 31	UNSPECIFIED 99
REMEMBERING 101																																
DISCRIMINATING 102																																
SORTING 103																																
IDENTIFYING 104																											●					
RECOGNIZING SEQUENCES 105																																
UNDERSTANDING 106																																
PROBLEM SOLVING 107																																
CRITICAL THINKING 108																																
CREATING, INVENTING 109																																
UNSPECIFIED 199																																
ATTENDING 201																																
FEELING FREE, NOT FEARING 202							●													●												
ENJOYING, LIKING 203																				●												
EXPECTING SUCCESS 204							●													●												
BELIEVING, VALUING 205																																
REFRAINING, RESPECTING 206											●	●															●					
WILLING TO DO 207																				●												
BEING CURIOUS 208																																
INITIATING 209																																
EMPATHIZING 210																																
CHECKUPS 301																																
PHYSICAL STATUS 302																											●					
PHYSICAL SKILLS 303			●																	●	●											
ANALYZING 401																																
PLANNING 402																																
MANAGING 403																																
COPING/PROBLEMS 404																																
COPING/FEELINGS 405																																

Figure 25

Process and Content of Characteristics Identified in Dimension 75: FINE MOTOR unvalidated

CONTENT

PROCESS

	01 VISUAL	02 AUDITORY	03 HAPTIC	04 SMELL	05 TASTE	06 NON-SPECIFIC	07 SELF	08 FAMILY	09 PEERS	10 ADULTS	11 SCHOOL	12 COMMUNITY	13 ORAL LANGUAGE	14 WRITTEN LANGUAGE	15 2ND LANGUAGE	16 NUMBERS	17 WTS. & MEASURES	18 MONEY	19 TIME	20 BODY MOVEMENT	21 MUSIC	22 VISUAL ARTS	23 DRAMATIC ARTS	24 POETRY & LITERATURE	25 EARTH/UNIVERSE	26 PLANTS, ANIMALS	27 HEALTH, NUTRITION	28 BUSINESS & MANUF.	29 BUILDING	30 TRANSPORTATION	31 SOCIETY	99 UNSPECIFIED
REMEMBERING 101																																
DISCRIMINATING 102																																
SORTING 103																																
IDENTIFYING 104														●		●																
RECOGNIZING SEQUENCES 105																																
UNDERSTANDING 106																																
PROBLEM SOLVING 107																																
CRITICAL THINKING 108																																
CREATING, INVENTING 109																																
UNSPECIFIED 199																																
ATTENDING 201																																
FEELING FREE, NOT FEARING 202																																
ENJOYING, LIKING 203																																
EXPECTING SUCCESS 204																																
BELIEVING, VALUING 205																																
REFRAINING, RESPECTING 206																											●					
WILLING TO DO 207																																
BEING CURIOUS 208																																
INITIATING 209																																
EMPATHIZING 210																																
CHECKUPS 301																																
PHYSICAL STATUS 302																											●					
PHYSICAL SKILLS 303			●											●						●	●	●										
ANALYZING 401																																
PLANNING 402																																
MANAGING 403																																
COPING/PROBLEMS 404																																
COPING/FEELINGS 405																																

Applied Strategies

What is posited as the Domain of Applied Strategies embraces many characteristics included in the other Domains, but the focus is on their appropriate use in different contexts. It is important not only that a child know or be able to do something, but also that he/she adept such knowledge or ability to the different requirements of varying concrete situations. This implies the competency to perceive and analyze relevant factors in the context of the overall environment, to select actions from one's repertoire of behaviors that are appropriate to the situation, and to carry out those actions while regulating one's emotions towards the attainment of an intended objective.

As illustrated in Figure 3 (p. 56) such a competency uses knowledges and skills, attitudes and habits, and physiological statuses from the Cognitive, the Social-Emotional and the Health and Physical Domains, and "executes" them (i.e., puts them to work together) in dealing with multi-faceted, real-life situations. This ability to organize one's other competencies into an integrated, purposeful course of action exemplifies what are termed "executive skills". These skills, as noted in Figure 3, are, by definition, unqiue to the Applied Strategies Domain.

Bronson (1974) uses the concept "executive skills" with precisely the same meaning as this project. She defines such skills as "the preschool child's functional level of competence in social interaction and in coping tasks ... skill in recognizing or identifying cues in a situation, skill in predicting possible sequences of events and outcomes of a situation, and skill in controlling both the self and the social or material 'other' in situation in order to effectively reach chosen goals".

Thus, Applied Strategies are complex abilities in which a child draws upon his or her competencies in any or all other domains of development and applies them in real-life situations. They are exemplified by executive skills, no matter what the context in which they are demonstrated. Illustrative are such abilities as carrying a task through to completion in the face of distractions, coping with an irate parent, using appropriately different forms of address and demeanor with peer as compared with adults, coping with racist behavior directed at oneself or a friend, devising a strategy for success in a physical or intellectual game, dealing with a teacher who appears to underestimate one's competence, handling a situation where physical combat is threatened, seeking appropriate aid when ill or injured, winning permission to "go out and play" from a reluctant parent, setting priority goals in the face of attractive alternatives, etc. The need for such abilities abound -- in home, school and community; they are manifested, or need to be manifested, as "the child's everyday effectiveness in dealing with his environment and later repsonsibilities in school and life" (OCD, 1975b, p. 1).

Inasmuch as the Applied Strategies Domain encompasses all other characteristics in all Domains plus executive skills, it is possible to conceptualize a very large number of constructs within the Domain. Such a structure would be counterproductive to the objectives of this project. Only two broad Dimensions are defined for measurement -- one involving Interpersonal Competencies and the other Task Competencies. The relationship between development in these two Dimensions and that in the Dimensions of the other Domains, and the need for more narrowly defined subconstructs within Applied Strategies, may be evaluated through analysis of data obtained in the use of the system of measures to be developed.

It is also recommended that the developer of measures for the Domain of Applied Strategies concentrate on the measurement of executive skills. The other types of characteristics will be the focus of the developers of measures for the other Domains; as such their inclusion in the Applied Strategies development efforts would represent a duplication of effort, and not the best use of the limited resources that will be available.

The executive skills identified by the participants in the input workshops for the Interpersonal and Task Competency Dimensions are listed, respectively, on pages 80 and 81 in Appendix D. They are also summarized in Figures 26, 27 and 28 by their process-content properties.

Rating of Characteristics

Although all of the characteristics identified as candidates for measurement stem from a common source -- the input workshops -- individual characteristics need to be further differentiated in at least two respects. Some may be considered more or less important by the participants and some may be considered particularly relevant to children of different ages. Data on these two attributes -- importance and age -- were obtained on the second day of the input workshops, as previously described (pp. 31-32). An analysis of these data yields information to further guide measures developers in selecting particular sets of characteristics as determinants of program effects. A more complete description of the input workshops procedures and of the data obtained is provided in A System for Selecting Characteristics to Measure Head Start Program Effects (Mediax, 1980).

It should be noted again that the participants in the input workshops, i.e., those responding to the questionnaires on which this analysis is based, did not constitute a representative or random sample of the greater population of Head Start parents, staff and K-2 public school teachers. Also, because questions asked on the second day were only those generated by the particular workshops on the first day, there is a different proportion of responses to the questions and not all questions are responded to by all demographic groups. Some of the conclusions based on these data in the following, therefore, may not be applicable to a more representative population.

Ratings of Importance

Overall, 79.5% of the 1,516 characteristics were rated by all the participants as "highly significant" or "significant." The overall ratings, on a four level scale were:

Highly significant	24.3%
Significant	55.2%
Neutral	17.9%
Negative	2.6%
	100.0%

From the above, it appears reasonable to conclude that the overall population of participants generally rated the characteristics as important to be measured. Appendix F includes a description of the method of calculating these overall ratings and lists the ratings of each characteristic by various categories of participant. There appear to be significant differences in the ratings of particular characteristics and in the ratings by particular categories of participant.

Figure 26

Process and Content of Characteristics Identified in the APPLIED STRATEGIES Domain unvalidated

CONTENT

PROCESS	VISUAL 01	AUDITORY 02	HAPTIC 03	SMELL 04	TASTE 05	NON-SPECIFIC 06	SELF 07	FAMILY 08	PEERS 09	ADULTS 10	SCHOOL 11	COMMUNITY 12	ORAL LANGUAGE 13	WRITTEN LANGUAGE 14	2ND LANGUAGE 15	NUMBERS 16	WTS. & MEASURES 17	MONEY 18	TIME 19	BODY MOVEMENT 20	MUSIC 21	VISUAL ARTS 22	DRAMATIC ARTS 23	POETRY & LITERATURE 24	EARTH/UNIVERSE 25	PLANTS, ANIMALS 26	HEALTH, NUTRITION 27	BUSINESS & MANUF. 28	BUILDING 29	TRANSPORTATION 30	SOCIETY 31	UNSPECIFIED 99
REMEMBERING 101																																
DISCRIMINATING 102																																
SORTING 103																																
IDENTIFYING 104																																
RECOGNIZING SEQUENCES 105																																
UNDERSTANDING 106																																
PROBLEM SOLVING 107																																
CRITICAL THINKING 108																																
CREATING, INVENTING 109																																
UNSPECIFIED 199																																
ATTENDING 201																																
FEELING FREE, NOT FEARING 202																																
ENJOYING, LIKING 203																																
EXPECTING SUCCESS 204																																
BELIEVING, VALUING 205																																
REFRAINING, RESPECTING 206																																
WILLING TO DO 207																																
BEING CURIOUS 208																																
INITIATING 209																																
EMPATHIZING 210																																
CHECKUPS 301																																
PHYSICAL STATUS 302																																
PHYSICAL SKILLS 303																																
ANALYZING 401							●	●																								●
PLANNING 402							●		●																							●
MANAGING 403									●	●	●		●						●				●									●
COPING/PROBLEMS 404								●	●	●		●																				●
COPING/FEELINGS 405							●		●																							●

Figure 27

Process and Content of Characteristics Identified in Dimension 81: TASK COMPETENCIES unvalidated

CONTENT

PROCESS	VISUAL 01	AUDITORY 02	HAPTIC 03	SMELL 04	TASTE 05	NON-SPECIFIC 06	SELF 07	FAMILY 08	PEERS 09	ADULTS 10	SCHOOL 11	COMMUNITY 12	ORAL LANGUAGE 13	WRITTEN LANGUAGE 14	2ND LANGUAGE 15	NUMBERS 16	WTS. & MEASURES 17	MONEY 18	TIME 19	BODY MOVEMENT 20	MUSIC 21	VISUAL ARTS 22	DRAMATIC ARTS 23	POETRY & LITERATURE 24	EARTH/UNIVERSE 25	PLANTS, ANIMALS 26	HEALTH, NUTRITION 27	BUSINESS & MANUF. 28	BUILDING 29	TRANSPORTATION 30	SOCIETY 31	UNSPECIFIED 99
REMEMBERING 101																																
DISCRIMINATING 102																																
SORTING 103																																
IDENTIFYING 104																																
RECOGNIZING SEQUENCES 105																																
UNDERSTANDING 106																																
PROBLEM SOLVING 107																																
CRITICAL THINKING 108																																
CREATING, INVENTING 109																																
UNSPECIFIED 199																																
ATTENDING 201																																
FEELING FREE, NOT FEARING 202																																
ENJOYING, LIKING 203																																
EXPECTING SUCCESS 204																																
BELIEVING, VALUING 205																																
REFRAINING, RESPECTING 206																																
WILLING TO DO 207																																
BEING CURIOUS 208																																
INITIATING 209																																
EMPATHIZING 210																																
CHECKUPS 301																																
PHYSICAL STATUS 302																																
PHYSICAL SKILLS 303																																
ANALYZING 401							●																									●
PLANNING 402							●																									●
MANAGING 403											●								●				●									●
COPING/PROBLEMS 404										●																						●
COPING/FEELINGS 405							●																									●

Figure 28

Process and Content of Characteristics
Identified in Dimension 82: INTERPERSONAL COMPETENCIES
unvalidated

CONTENT

PROCESS	VISUAL 01	AUDITORY 02	HAPTIC 03	SMELL 04	TASTE 05	NON-SPECIFIC 06	SELF 07	FAMILY 08	PEERS 09	ADULTS 10	SCHOOL 11	COMMUNITY 12	ORAL LANGUAGE 13	WRITTEN LANGUAGE 14	2ND LANGUAGE 15	NUMBERS 16	WTS. & MEASURES 17	MONEY 18	TIME 19	BODY MOVEMENT 20	MUSIC 21	VISUAL ARTS 22	DRAMATIC ARTS 23	POETRY & LITERATURE 24	EARTH/UNIVERSE 25	PLANTS, ANIMALS 26	HEALTH, NUTRITION 27	BUSINESS & MANUF. 28	BUILDING 29	TRANSPORTATION 30	SOCIETY 31	UNSPECIFIED 99
REMEMBERING 101																																
DISCRIMINATING 102																																
SORTING 103																																
IDENTIFYING 104																																
RECOGNIZING SEQUENCES 105																																
UNDERSTANDING 106																																
PROBLEM SOLVING 107																																
CRITICAL THINKING 108																																
CREATING, INVENTING 109																																
UNSPECIFIED 199																																
ATTENDING 201																																
FEELING FREE, NOT FEARING 202																																
ENJOYING, LIKING 203																																
EXPECTING SUCCESS 204																																
BELIEVING, VALUING 205																																
REFRAINING, RESPECTING 206																																
WILLING TO DO 207																																
BEING CURIOUS 208																																
INITIATING 209																																
EMPATHIZING 210																																
CHECKUPS 301																																
PHYSICAL STATUS 302																																
PHYSICAL SKILLS 303																																
ANALYZING 401								●																								●
PLANNING 402									●																							
MANAGING 403									●	●			●																			●
COPING/PROBLEMS 404								●	●	●		●																				
COPING/FEELINGS 405							●		●																							

Most positively rated characteristics. For each characteristic rated in the input workshops, the mean of the ratings made by all participants was calculated. This mean was then compared, using a t test for unmatched groups, to the distribution of mean ratings averaged across all other characteristics. This procedure resulted in a value of t associated with each characteristic that represents the statistical significance of the difference between the rating of that characteristic and the average rating of all other characteristics; the larger the value of t, the more positively that item was judged. Although this application of the t test does not, strictly speaking, meet the assumption of independence, it does suggest a minimum criterion for judging the magnitude of these mean differences. The 50 characteristics whose ratings are most different from the average, in a positive direction, are listed in Table VIII.

The extent to which each content category is represented among the 50 most positively rated characteristics was examined. Table IX shows the number of characteristics in each content area appearing among the most positive 50, along with the percentage (e.g., the Visual content area is represented by one characteristic among the most positive 50; thus, Visual content represents 2.0% of the most positive characteristics compared with its representation of 2.8% among all characteristics). Two content areas stand out as being much more highly represented in the list of most positive -- "27. Health" with 3.5 times as large a representation and "31. Society" with 3.6 times as many characteristics in the most positive 50 as in the total.

Inspection of the specific characteristics under these content areas shows they are closely related. Of the 22 characteristics under "27. Health", 15, including 9 out of 10 of the most highly rated, deal with safety. Also, all 4 characteristics under "31. Society" refer to rules of society concerned with safety. These two content categories related to physical well being, account for 52% of the list of the most positively rated although they only account for 14.8% of the characteristics on the overall list.

It therefore seems reasonable to conclude that characteristics associated with basic physical health and safety are very much more highly rated than characteristics in any other area.

A second highly rated content area, somewhat overlapping health and safety, involves what are commonly referred to as basic skills. Included, additionally, here are those characteristics on the list concerned with basic colors, reading and writing one's own name, math skills, knowing the alphabet and reading readiness skills.

Most negatively rated characteristics. Applying the same procedures as for the most highly rated characteristics, a list of the 50 lowest rated characteristics was compiled. These are listed in Table X. Their representation is shown in Table IX. It should be emphasized that these "negative" ratings are relative. They represent characteristics of development that are seen as less important by the workshop participants. As can be seen in Table X (column 2), the mean rating of these items is often at the level of neutral (2.0) or above.

The most frequently represented category of content in the lowest rated 50 was "09. Peers". However, its representation (14.0%) was not substantially greater than in the overall list (10.1%). The content area that stands out for its low rating is "15. Second or Other Language." It occurs six times in the list of most negative 50, or 12%, compared to its representation in the overall list of characteristics of only 1.3%. It should be noted that this content area is highly differentiated in importance ratings by particular categories of participants; therefore, a conclusion of a

TABLE VIII

50 Characteristics Rated Most Positive by Participants in Input Workshops

t Value (Positive)	Mean	Group Size	Path Number	Statement of Characteristic
15.8	3.6	359	206.27.019	Do not play with fire or flammable objects, hot objects,....
14.6	3.5	361	106.13.011	Express simple needs; thoughts & feelings; be able to comm. ....
14.6	3.6	296	206.27.046	Not tamper with electrical outlets, appliances
14.4	3.5	343	206.27.018	Do not play in street; play only in safe areas
14.3	3.5	361	206.27.020	Do not play with knives or other sharp or dangerous objects....
13.3	3.6	229	104.07.006	Know their own name
12.3	3.5	314	202.07.013	Be able to go to the bathroom without assistance
12.1	3.5	280	207.27.012	Have good safety habits
12.0	3.5	256	203.07.007	Feel good about themselves; have a good self image
11.7	3.4	345	106.31.008	Recognize traffic lights and other safety signs and know....
11.4	3.5	236	104.27.032	Recognize poisonous labels and symbols
10.8	3.4	359	105.27.001	Can cross street on way to school; obey traffic rules;....
10.6	3.6	156	206.27.066	Watch for cars; know when to get out of street; not play in street
10.4	3.5	238	206.27.030	Keep dirty fingers and objects out of mouth; do not....
10.3	3.4	359	301.27.004	Have initial dental examinations; have first trip to dentist
10.2	3.6	156	105.27.003	Know how to avoid hot stove and other dangerous objects in....
10.2	3.4	279	202.08.006	Feel free to say they are frightened; to confide in parents;....
10.2	3.4	314	404.10.002	Seek help when in distress; be able to react in emergencies;....
10.1	3.4	344	206.27.005	Avoid doing things that endanger or hurt others
9.8	3.3	363	104.01.005	Recognize basic colors
9.8	3.3	344	104.14.008	Be able to write name
9.7	3.4	228	207.27.011	Have good health habits
9.6	3.4	227	104.31.006	Safety hazards; know playground rules and equipment
9.6	3.3	356	202.07.026	Feel free to cry when frustrated or hurt; feel free to respond....
9.2	3.4	243	199.16.006	Have number and arithmetic skills
9.2	3.4	279	206.27.060	Use good toilet habits; go to bathroom w/o assistance
9.1	3.4	213	206.27.017	Do not eat from floor or ground; do not taste things out of....
9.0	3.3	358	104.14.014	Distinguish own name from others
9.0	3.4	264	201.13.002	Have ability to listen and focus attention; be able to listen
8.8	3.4	241	206.11.009	Follows thru when given directions; follows directions from....
8.8	3.4	274	104.07.005	Know own address and important phone numbers
8.7	3.3	358	202.08.002	Be able to join a group of children & let go of parents; ....
8.7	3.4	246	104.08.007	Know names of their parents, sisters, brothers
8.7	3.3	306	206.10.010	Do not speak to, take things from, go with strangers
8.6	3.3	360	202.07.017	Be free to express any and every emotion; to show frustration,....
8.6	3.3	351	206.31.004	Not take things that don't belong to them; respect others'....
8.6	3.3	349	202.09.010	Be able to work and play in small groups and get along well ....
8.6	3.3	361	104.14.028	Recognize letter names in alphabets; read abc's (lower case).....
8.4	3.3	315	206.27.028	Go to bed when tired without being told; getting proper rest
8.4	3.3	357	302.27.015	Have eye-hand coordination; develop fine motor skills
8.4	3.5	132	104.27.011	Know not to take medicine on their own
8.3	3.3	287	205.12.002	Know that everyone is equal; that there are no superor or....
8.3	3.3	289	202.07.023	Describe their thoughts and feelings to others; be able to express....
8.2	3.3	330	206.31.010	Understand and observe simple rules of society....
8.2	3.4	192	206.27.043	Not climb or play near open windows or in high places
8.2	3.4	206	207.27.009	Have good dental habits
8.2	3.4	203	104.14.017	Have reading readiness skills
8.2	3.3	363	206.27.002	Apply personal cleanliness; physical hygiene; bathe wash brush....
8.1	3.4	192	301.27.003	Have general medical check-ups regularly

TABLE IX

COMPARISON OF IMPORTANCE RATINGS OF SPECIFIC CONTENT CHARACTERISTICS

Content	Representation In Overall List		Representation In Most Positive 50			Representation In Most Negative 50			Representation In Most Neutral 50		
	No.	%	No.	%	f*	No.	%	f*	No.	%	f*
01. Visual	43	2.8	1	2.0	.7	0	—	—	0	—	—
02. Auditory	9	.6	0	—	—	0	—	—	0	—	—
03. Haptic	44	2.9	0	—	—	1	2.0	.7	5	10.0	3.4
04. Smell	6	.4	0	—	—	0	—	—	0	—	—
05. Taste	4	.3	0	—	—	0	—	—	0	—	—
06. Senses, Non-Specific	12	.8	0	—	—	0	—	—	0	—	—
.07 Self	126	8.3	8	16.0	1.9	5	10.0	1.2	1	2.0	.2
08. Family And Own Culture	98	6.5	3	6.0	.9	2	4.0	.6	7	14.0	2.2
09. Peers	153	10.1	1	2.0	.2	7	14.0	1.4	1	2.0	.2
10. Adults	95	6.3	2	4.0	.6	4	8.0	1.3	2	4.0	.6
11. School	80	5.3	1	2.0	.4	5	10.0	1.9	1	2.0	.4
12. Community And Other Cultures	54	3.6	1	2.0	.6	1	2.0	.6	2	4.0	1.1
13. Oral Language	78	5.1	2	4.0	.8	4	8.0	1.6	6	12.0	2.4
14. Written Language	92	6.1	4	8.0	1.3	1	2.0	.3	1	2.0	.3
15. Second Or Other Language	20	1.3	0	—	—	6	12.0	7.7	2	4.0	3.0
16. Numbers	54	3.6	1	2.0	.6	2	4.0	1.1	0	—	—
17. Weights And Measures	4	.3	0	—	—	0	—	—	0	—	—
18. Money	8	.5	0	—	—	0	—	—	0	—	—
19. Time	14	.9	0	—	—	0	—	—	0	—	—
20. Sports; Dance	65	4.3	0	—	—	6	12.0	2.8	10	20.0	4.7
21. Music	13	.9	0	—	—	0	—	—	2	4.0	4.4
22. Visual Arts	21	1.4	0	—	—	0	—	—	2	4.0	2.9
23. Dramatic Arts	17	1.1	0	—	—	2	4.0	3.6	1	2.0	1.8
24. Poetry And Literature	14	.9	0	—	—	0	—	—	0	—	—
25. Earth; Universe; Physical Science	21	1.4	0	—	—	0	—	—	0	—	—
26. Plants; Animals; Physiology	66	4.4	0	—	—	1	2.0	.5	0	—	—
27. Health; Nutrition; Dental Health	191	12.6	22	44.0	3.5	1	2.0	.2	4	8.0	.6
28. Business; Manufacturing	18	1.2	0	—	—	0	—	—	0	—	—
29. Building And Construction	8	.5	0	—	—	0	—	—	0	—	—
30. Transportation	12	.8	0	—	—	0	—	—	2	4.0	5.0
31. Society; Social Order	33	2.2	4	8.0	3.6	2	4.0	1.8	1	2.0	.9
99. Unspecified	43	2.8	0	—	—	0	—	—	0	—	—

$$^*f = \frac{\%\text{ in 50}}{\%\text{ in overall}}$$

TABLE X

50 Characteristics Rated Most Negative by Participants in Input Workshops

t Value (Negative)	Mean	Group Size	Path Number	Statement of Characteristic
18.1	1.9	163	104.14.009	Be differentiated as (knowing that you are) good or bad readers
18.1	1.6	233	199.16.003	Be good with numbers; boys better than girls
18.1	1.5	176	199.16.005	Be good with numbers; girls better than boys
18.1	1.9	237	202.07.031	Have temper tantrums when they are upset; feel free to throw...
18.1	2.0	215	202.09.031	Should not compete with other children (not outdo others)
18.1	2.0	229	202.11.009	Talk and move about in the classroom at all times
18.1	1.5	122	205.09.003	Believe that boys are superior to girls
18.1	1.5	122	205.09.055	Believe that girls (women) are superior to boys
18.1	1.6	226	210.08.006	Lie to protect brothers or sisters from spanking
18.1	1.6	226	210.09.022	Lie to protect close friends from punishment
17.9	2.3	345	210.10.002	Cry if they see someone hurt or sad
16.3	2.1	193	202.07.035	Show bad facial expressions (e.g., stick tongue out)
16.2	1.7	86	206.15.001	Speak only English
16.0	1.9	125	206.13.013	Use only standard English regardless of language of their home
16.0	1.9	125	206.15.002	Use only standard English (in the classroom) regardless....
15.7	1.2	44	108.07.003	Use sickness as a copout for wanting to rest/not going to school
15.4	2.0	121	104.15.006	Read words in non-English language before English
15.1	1.3	44	403.10.001	Be able to manipulate parents and teachers
15.0	2.2	213	106.31.013	Understand that success is often based on who you are....
14.7	2.3	226	207.20.004	Should strive to excel at sports and games
14.4	2.1	139	199.11.005	Should be at similar level of development in all regards....
14.3	1.8	80	202.07.020	Children who are boys/girls do things boys/girls do
14.3	1.6	56	202.11.006	Show they are ready for school by feeling free to get....
14.3	2.0	127	202.13.005	Correct the incorrect speech of adults
13.8	2.4	285	203.20.001	Enjoy dancing (social)
13.7	2.3	243	303.20.021	Be able to skate, use skooter skates
13.6	2.1	120	105.26.014	Watch (know how) a child being born
13.3	2.3	184	204.07.007	Take risks for good causes
13.2	2.3	190	202.15.004	Speak standard English in the classroom and native language....
13.1	1.8	67	207.11.009	Say they want to be teacher when...grow up, react positively....
13.0	2.4	241	106.23.008	Play Superman, Spiderman or Superwoman, TV characters
13.0	2.5	351	210.09.008	Be able to feel compassion or sorrow for others; show....
12.9	2.1	128	202.13.004	Correct speech of their parents
12.8	2.5	340	202.11.007	Speak their dialects in classroom; be able to pronounce words....
12.8	2.5	340	202.15.005	Speak their dialects in classroom; pronounce words correctly
12.5	2.3	212	106.31.005	Know that culture influences expectations for boys/girls
12.0	2.5	300	106.13.005	Be able to re-tell a 10 minute story in their own words
11.9	2.3	186	303.20.026	Be able to turn somersaults and do flips
11.8	2.3	153	303.20.003	Be able to climb trees
11.5	2.4	248	104.15.001	Be able to speak and understand a second language other than....
10.9	2.5	270	201.10.003	Be able to listen to others while moving about
10.6	2.6	354	210.10.001	Be able to feel compassion or sorrow for others, show....
10.6	1.9	51	303.03.026	Eat with chopsticks
10.5	2.4	193	104.12.013	Understand significance of special holidays (Mother's Day, etc.)
10.4	1.7	38	201.09.003	Close out the world when in deep conversation with another child
10.4	2.1	82	206.27.056	Stop cars by raising their hand when crossing the street
10.3	2.3	146	210.08.005	Feel sorry if parent expresses anger
10.2	1.8	41	303.20.037	Pick up weights
10.2	2.4	179	202.09.014	Compete in group rather than individually
10.0	2.3	105	106.23.001	Be able to take-off on movies or TV

generalized rating would be erroneous. This content area provides an example of the diverse values as discussed in Chapter III.

The next most significantly represented content category in the list of negative characteristics is "20. Sports and Dance." It appears in the list 2.8 times more than would be predicted by its representation in the overall list. This suggests that theparticipants did not want Head Start to become involved with teaching at least some sports and dance activities.

Inspection of the list of more negative characteristics also reveals obviously controversial items that were understandably rated negatively. Their appearance provides assurances on the validity of the methodology and the data.

Characteristics with the most neutral ratings. The four-level scale used to interpret importance ratings of characteristics included a neutral level, which introduces considerations different than the positive or negative ratings. A positive rating implies the respondent feels it is important for the child to acquire the characteristic. A negative rating is interpreted as implying that the repsondent feels it is undesirable for the child to acquire the characteristic or that it is undesirable for Head Start and the child to devote energies to its acquisition. A neutral rating, however, can be interpreted as either valuing the acquisition of the characteristic but not necessarily during Head Start, or as not particularly valuing whether the characteristic is acquired or not.

The percent of responses that were neutral were calculated for each characteristic. Overall, for all characteristics, 17.9% of the responses were neutral. The percentage of neutral responses for the 50 characteristics most neutrally rated ranged beteen 44.8% and 64.2% (see Table XI). Analysis of the content of these characteristics provides further insights to guide test developers.

The two characteristics most frequently rated neutral were both in content category "20. Sports, Dance." Also, 4 out of the top 10 on the list of 50 were of this content category. This appears to reflect a general apathy about Head Start becoming involved in the development of children's sports or dancing skills; a conclusion previously arrived at on the basis of the analysis of the characteristics rated most negative.

Similarly, overall the participants appear unconcerned about young children "rushing" to gain more advanced understandings and skills. This was also communicated in the wording of some characteristics that referred to "children being ready" or "at their best time". In the list of most neutrally rated characteristics are included "understanding that racism is associated with economic status", "be able to draw what they see", "learn how to make bridges", "be able to sew", "help shop for food", "learn about medical discoveries", "read with speed", "learn about statutes", "learn breathing and relaxation techniques", "understand eye language" "be able to sing", "learn how airplanes are made", and "understand TV". Unlike the sports and dance content area, which was frequently represented in the list of most negatively rated characteristics, none of these more advanced learnings are referred to among the characteristics rated negative. We interpret this as a desire for the acquisition of these skills, but not necessary during the ages in which their children are affected by Head Start.

Demographic-group differences in importance ratings. An important hypothesis influencing many of the recommendations to measures developers is that there is diversity in the paths and indicators of Head Start children's development. Such

TABLE XI

50 Characteristics Most Rated Neutral by Participants in Input-Workshops

Neutral Rating %	Group Size	Path Number	Statement of Characteristics
64	81	303.20.046.74	Be able to wrestle
63	82	303.20.027.74	Be able to use pogo sticks
60	74	203.12.002.68	Enjoy collecting lucky money on Chinese New Years
57	121	104.14.006.53	Read words in non-English language before English
57	80	206.27.053.72	Should take fluoride tablets
56	74	104.20.002.56	Learn folk dancing
56	78	404.10.001.82	Compromise with adults who may not like them
55	70	106.31.012.56	Understand that racism is related to economic status
55	49	303.03.023.75	Braid hair
55	49	303.20.014.75	Be able to pitch horseshoes
55	45	303.22.004.75	Be able to draw what they see
54	55	203.20.002.62	Enjoy dancing to rhythm of music; salty spiced
53	92	210.09.027.66	Want Mom to buy for friends also
51	54	104.30.006.55	Learn how they make missiles
51	58	207.08.003.68	Help shop for food
51	39	303.03.014.75	Be able to punch
51	80	303.03.016.75	Be able to sew
50	44	104.27.023.71	Learn about medical discoveries
50	80	106.13.010.52	Correct the incorrect speech of other children their own age
50	50	202.08.007.64	Learn from brothers and sisters
50	80	210.08.001.61	Ask mothers not to spank brothers or sisters
50	146	210.08.005.61	Feel sorry if parent expresses anger
50	108	303.20.040.74	Play hopscotch
50	36	303.22.009.75	Can make -monitos- (pictures)
49	135	104.13.018.52	Understand other children's dialects and languages
49	59	302.27.007.74	Be able to touch toes
49	51	303.03.026.75	Eat with chopsticks
48	189	104.21.002.52	Can sing a song after they have heard it a few times
48	127	202.13.005.68	Correct the incorrect speech of adults
48	52	202.15.001.63	Be bilingual; bicultural; bilateral (feeling comfortable)
48	54	204.08.007.68	State enjoyment in helping with household tasks;....
48	226	207.20.004.74	Should strive to excel at sports and games
48	345	210.10.002.61	Cry if they see someone hurt or sad
48	50	303.03.017.75	Be able to shell beans and peas
47	46	104.12.009.56	Learn about statues
47	55	104.27.005.74	Know breathing and relaxation techniques
47	53	106.13.016.52	Understand eye language
47	48	106.14.016.53	Read with speed (with understanding)
47	139	199.11.005.64	Should be at similar level of development in all regards....
47	128	202.13.004.68	Correct speech of their parents
47	48	303.20.035.74	Know how to use jump board
46	49	104.21.001.52	Be able to sing
46	54	105.30.001.55	Learn how airplanes are made
46	78	106.13.009.52	Be able to understand and explain talk of 3-yr. old
46	94	203.23.001.64	Enjoy watching TV programs
46	182	207.08.002.68	Ask "what else" after completing task requested by mother
36	135	210.08.00761	Pamper a crying parent
46	243	303.20.021.74	Be able to skate, use skooter skates
46	186	303.20.026.74	Be able to turn somersaults; be able to do flips
45	44	102.07001.51	Imitate things that happened to them

diversity should therefore be evidenced in the importance ratings of characteristics identified in the input workshops.

Although the analysis of data on importance is confounded by the factors previously described (p. 99) there is clear evidence of such differences among the various demographic groups participating. There is also evidence of two other kinds of diversity among the demographic groups that affects importance ratings -- (1) in their methods of expression of importance ratings, and (2) in the homogeneity of their values.

These determinations were derived from two different approaches to the analysis of the day two data, as summarized for each characteristic by demographic group in Appendix F. First, the mean rating of each characteristic was calculated for each demographic group and compared with the mean rating of that characteristic by all other participants. In computing means, weights were assigned each category or response as follows:

Highly Significant - 4
Significant - 3
Neutral - 2
Negative - 1

For each characteristic, the mean rating by a demographic group is considered different from the mean of all other participants if the means were significantly different ($p<.01$) and if the difference between the means was .5 or greater. A list of each characteristic so differentiated by each demographic group is included in Section IV of Appendix F. A summary of the number of positively and negatively differentiated characteristics is presented in Table XII. The table shows that, for example, Black workshop participants rated 18 characteristics as more important (i.e., more positive) than did the other participants. These 18 characteristics were rated by an average of 24.3 Black participants, and the number of Black participants rating each characteristic ranged from 6 to 62. There were 4 characteristics which Blacks rated as more negative than the average participant, and these ratings were based on an average of 16.3 Black participants.

Second, a compilation was prepared of characteristics identified by more than 50% of a particular ethnic group as "highly significant", "neutral", and "negative", and where such rating was more than 20 percentage points greater than the response of all participants. The 50% requirement was used to select characteristics on which there was a clear concensus within each demographic group. The greater-than-20% requirement was imposed to further isolate for each demographic group ratings that were clearly differentiated form those of the overall participants. The characteristics so identified for each demographic group are listed in Section V of Appendix F. A summary of the number of characteristics so differentiated by each demographic group is presented in Table XIII.

As previously explained, although many of the characteristics were rated by all or most of the participants, some characteristics were only rated in one or a few of the input workshops. There would be a tendency for such characteristics to appear disproportionately differentiated as there would be more confounding of the demographic groupings by regional factors, and the smaller numbers would increase the likelihood of chance differences. To permit accounting for such bias in the data the number of participants of the particular demographic group rating each characteristic is provided in the listings and tables.

TABLE XII

SUMMARY OF POSITIVELY & NEGATIVELY DIFFERENTIATED IMPORTANCE RATINGS OF DEMOGRAPHIC GROUPS (WITH 15 OR GREATER DIFFERENCE FROM MEAN OF REMAINDER OF POPULATION ON A SCALE OF 3.0 — RANGING FROM NEGATIVELY RATED [1.0] TO HIGHLY SIGNIFICANT [4.0])

		POSITIVELY DIFFERENTIATED CHARACTERISTICS				NEGATIVELY DIFFERENTIATED CHARACTERISTICS			
DEMOGRAPHIC GROUP	MAX N	RANGE OF DIFFERENCE IN MEANS	NUMBER DIFFERENTIATED	AVERAGE N	RANGE OF N's	RANGE OF DIFFERENCE IN MEANS	NUMBER DIFFERENTIATED	AVERAGE N	RANGE OF N's
Black	131	.5-.10	18	24.3	6-62	.5-.8	4	16.3	13-18
White	133	.5-.8	6	18.0	11-24	.5-.7	8	38.8	15-80
Hispanic	67	.5-.7	5	13.2	10-26	.5-.8	16	16.5	7-29
Native American	17	—	0	—	—	.5-1.6	6	7.0	3-11
Asian	16	.5-.9	10	7.3	7-8	.5-1.1	33	9.5	4-16
Parents	183	.5-.8	20	40.1	27-127	.5-.7	15	27.2	18-60
Head Start Staff	132	—	0	—	—	.5-.6	12	28.9	17-61
Public School Teachers	62	.5-.9	47	19.0	7-61	.5-1.0	7	16.6	9-24
Migrant	16	—	0	—	—	.5-15	21	13.4	2-16
Rural	126	.5-.7	3	33.7	9-54	.5-.7	3	13.3	13-14
Urban	222	.5-.7	4	37.3	19-61	.5-.7	2	15.5	14-17
Northeast	82	.5-1.0	16	79.3	73-82	.5-.8	3	79.7	77-82
Southeast	50	.6-.6	1	43	43	.5-.8	5	47.8	45-50
Central	93	.5-1.3	11	40.5	14-76	—	0	—	—
South	103	.8-.8	1	54	54	.5-1.0	9	58.6	37-101
West	36	—	0	—	—	.5-.6	17	33.1	31-35

TABLE XIII

SUMMARY OF CHARACTERISTICS CONSISTENTLY (>50%) RATED DIFFERENTIALLY FROM RATINGS OF OVERALL PARTICIPANTS (GREATER THAN 20% DIFFERENCE IN RATINGS)

*Number Of Characteristics Rated Differentially**

Demographic Group	*Max N*	*Highly Significant*		*Neutral*		*Negative*	
		#	*Aver. N*	*#*	*Aver. N*	*#*	*Aver. N*
Black	131	34	26.5	2	11.5	2	19.5
White	133	18	33.2	26	35.2	3	24
Hispanic	67	12	9.8	37	13.4	2	29
Native American	17	4	5.5	15	10.5	2	.5
Asian	16	16	7.3	45	8.2	1	6.0
Parents	183	18	20.4	2	22.5	1	25
Head Start Staff	132	10	9	11	18.2	4	13.8
Public School Teachers	62	157	26.8	30	13.8	6	10.7
Migrant	16	0	—	18	13.7	2	13
Rural	126	7	15.4	14	20.6	0	0
Urban	222	4	20.8	8	15.6	2	18.5
Northeast	82	9	79.8	12	79.9	0	0
Southeast	50	2	50.0	2	47.5	8	45
Central	93	18	51.7	5	33.6	0	2.0
Southwest	103	0	0	1	42.	6	59.3
West	36	0	0	8	33.9	1	34.0

* *1. N refers to the total number of participants in each demographic group; # refers to the number of characteristics rated.*

2. Only items are included in which the number of respondents in a demographic group is more than or equal to 5 and also the number of other respondents is equal to or more than 5.

The extensiveness of both these listings testifies to the existence of considerable diversity in the ratings by the various demographic groups, although to a larger degree in some than in others. All racial/ethic groupings distinctively value some characteristics. The Asian population appears to be more highly differentiated than the others. Parents and Head Start Staff as separate participant groupings by participant role, do not exhibit a group value; teachers on the other hand are the most highly differentiated group on any basis, particularly with respect to basic skills with academic significance (see Appendix F, Section V). The groupings by living style (migrant, rural, urban) exhibit only limited differentiations; those few are probably more a function of the influence of their respective disproportionate inclusion of other demographic groupings than of distinctive values of their own. The same influences appear in the differentiations of the various regional groupings, although the participants from theCentral (U.S.) had the most differentiated ratings, influenced by their smaller numbers and less representative populations.

The limitations of the data preclude traditional statistical procedures for analysis of variance; they also preclude generalizations on the nature of differences attributable to particular demographic groupings. The analysis that has been made permits recommending that measures development should recognize, in particular, the diversity in ethnic/racial groups and, where appropriate, the different emphases of Head Start parents and staff from public school teachers.

Age Ratings

Input-workshop participants were asked during the second day to indicate the age at which they thought children should attain each characteristic identified. It was recognized then, as it is now, that the participants had very questionable credentials for such a judgment; many had not had personal experiences with children at the age at which the particular characteristic was attained. Also,even for those participants who had such experiences, their perceptual and memory skills were untested. Nevertheless, their judgments were elicited to provide a rough guide on the appropriateness of the charcteristics and the dimensions to which they were assigned for measuring program effects on 3 to 7 year old children.

The mean age ratings of each characteristic hypothesized as associated with each dimension are listed in Appendix G. The ages are summarized in Table XIV.

A review of the data indicates that a large majority of the characteristics have been judged to be desirably attained between the ages 4 to 5 the primary Head Start year. This is not surprising considering that the participants were most familiar with Head Start children and had been asked in the workshop to focus on characteristics they desired Head Start to affect.

As the question on age was asked, an optimistic response could be anticipated. Therefore, it should be expected that, in fact, many of the characteristics will also be of use to measure program effects in later years than they were rated.

Also, if the standard deviations associated with the mean age for each characteristic (see listings, Appendix G) can be assumed to be indicative of a real difference among children in that age group, then characteristics can be assumed to be useful for measurement over a two to-three year span (approximately equal to plus and minus one standard deviation).

TABLE XIV

DISTRIBUTION OF MEAN AGE RATINGS BY INPUT-WORKSHOP PARTICIPANTS OF CHARACTERISTICS ASSOCIATED WITH EACH DIMENSION

		Number Of Characteristics			
Dimension	*Total No.*	*Mean Age ≤3.99*	*Mean Age 4.00-4.99*	*Mean Age 5.00-5.99*	*Mean Age ≥6.00*
.51 Perception	**75**	**25**	**49**	**1**	**0**
.52 Language	**96**	**26**	**62**	**7**	**1**
.53 Reading	**78**	**2**	**18**	**33**	**25**
.54 Math	**86**	**7**	**36**	**33**	**10**
.55Nature & Science	**103**	**11**	**65**	**24**	**3**
.56 Social Organization	**75**	**8**	**41**	**24**	**2**
.61 Sensitivity To Feelings Of Others	**70**	**29**	**39**	**2**	**0**
.62 Expression Of Own Feelings	**85**	**48**	**36**	**1**	**0**
.63 Self-Concept	**162**	**59**	**87**	**15**	**1**
.64 Attitudes Toward School	**177**	**49**	**99**	**24**	**5**
.65 Independence	**22**	**9**	**10**	**3**	**0**
.66 Sharing And Competing	**51**	**11**	**38**	**2**	**0**
.67 Peer Relationships	**148**	**48**	**96**	**4**	**0**
.68 Adult Relationships	**235**	**80**	**129**	**25**	**1**
.71 Health & Safety	**100**	**55**	**38**	**5**	**2**
.72 Dental	**48**	**12**	**31**	**5**	**0**
.73 Nutrition	**58**	**32**	**22**	**4**	**0**
.74 Gross Motor	**82**	**24**	**46**	**12**	**0**
.75 Fine Motor	**62**	**15**	**36**	**9**	**2**
.81 Task Competencies	**30** *a*	**1**	**21**	**7**	**1**
.82 Interpersonal Competencies	**22** *a*	**1**	**14**	**7**	**0**

a — Totals only include executive skills

Considering that diversity may be anticipated in the age ranges at which particular children attain particular characteristics and the other issues discussed above, it is recommended that objective statistical procedures be utilized by measures developers to establish hierarchical relationships in the attainment of characteristics. These procedures will also help identify the particular characteristics, out of the universe of those identified in the input workshops and those as may be added, that will be markers of increments of development along the various paths being followed by Head Start children. These markers are the critical characteristics for which measures of program effects must be developed.

V

TECHNICAL ISSUES FOR MEASURES DEVELOPMENT

Chapter II summarized the procedures used in the assessment of available instruments that have been used in Head Start evaluations. Chapter IV focused on the question: "What is to be measured?" The purpose of this chapter is to present and discuss the technical issues that are faced as measures developers proceed from an understanding of what is to be measured to the task of how to do the measuring.

There are major technical issues to be addressed in developing any measure or system of measures; these will obviously have to be dealt with in this project. In addition, however, this project faces special challenges as a function of the special purposes of these measures, the philosophical underpinnings of the project, and how the measures will be used in future evaluations (see especially the discussion in the Overview of the eight ways in which the new measures will differ from existing ones). Thus, this project strives to address programmatic concerns, not only in defining what is to be measured, but in resolving issues of how to measure these attainments.

After re-emphasizing the conclusions of Chapter II regarding the need for new measures, this chapter considers a number of technical issues, beginning with a discussion of issues in reliability and validity. Next, the chapter reviews some of the major analytic issues that will arise when data from these measures become available. In the third part of the chapter measurement procedures and techniques likely to be involved in future Head Start data collections are discussed in terms of how they relate to the proposed nature of the measures; certain requirements of the techniques of measurement naturally stem from the basic assumptions underlying the measures themselves.

Shortcomings of Existing Measures

As issues in assessing reliability and validity, issues in data analysis of Head Start measures, and data collection and measurement procedures are discussed in this chapter, it is worthwhile to keep in mind the state of the art that gives rise to these concerns. The final section of the chapter on methodology describes the procedures followed in this project in analyzing available measures. A review of the criteria used in that analysis and its finding that few existing measures are adequate for use in Head Start evaluations suggests issues that must be confronted in the development of new measures.

Criteria for Judging Measures

Five broad criteria were used in the initial screening of available measures:

1. Does the instrument attempt to measure national or local Head Start objectives?

2. Is it appropriate for Head Start/K-2 children's ages, ability levels, ethnicity, and bilingual status?

3. Is it practical for administration by paraprofessionals?

4. Has it been used in other major evaluation studies?

5. Has it been demonstrated to have good psychometric characteristics?

The measures that appeared to satisfy this screening were then examined on the basis of more specific criteria. The major criteria used in this appraisal, and the conclusions from the review, are as follows:

Relevance of content. The content of existing measures only partially matches Head Start program content. Most measures have been developed for other purposes, outside the Head Start context, so have only partial content validity for evaluating Head Start programs. In addition, there is very little evidence that measures are capable of detecting program-related growth.

Psychometric characteristics. Reasonable reliability indices are reported for most measures that have actually been used in Head Start evaluations. But most are based on internal consistency conceptions of reliability rather than test-retest reliability. Thus, if the content of the measures is questionable, the fact that the content is measured consistently within itself does not provide strong recommendation for using the measures in evaluations. Many evaluations, if they examine the question at all, also find fairly substantial correlations between measures that are designed to assess different dimensions of program effects.

Although research relating to the relationships between scores on measures used in evaluations and other measures of the same construct have been carried out, very few measures have any evidence of construct validity. None possesses anything approaching adequate construct validation.

Some criterion-related validity studies have been conducted in the context of program evaluations, in addition to studies reported by the test publishers. Criterion-related validity is claimed in relation to a number of concurrent and predictive criteria. The coefficients are not impressively high, however, and the criterion measures frequently bear only tangential relationships to program criteria relevant to Head Start.

Racial, ethnic, cultural and SES bias are often discussed by measures developers and users, and judgments are sometimes cited to the effect that there is "no apparent" bias. Very few studies report systematic research into this question, however.

Related to this issue is the question of norms. The majority of measures used have not been norm referenced; but the norm-referenced tests that have been used have had inadequate standardization in terms of the representativeness of their samples to the Head Start population. The use of norm-referenced tests has been justified on the grounds that program effects will be judged by raw-score comparisons with a non-treatment group and not by comparison to the norms.

Practical considerations. An important accomplishment of previous Head Start evaluations is the development of systematic procedures (including test formats, selection of data collectors, training systems, and quality control monitoring) that make large-scale assessment of program effects on children practical. These practical features should be retained and, if possible, improved upon in the measures to be developed in this project. (Unfortunately, practical considerations have often outweighed theoretical and technical considerations, and many evaluation reports have pointed to this dilemma as an issue to be addressed in designing new evaluations and the measures to be used in them.)

Judgments of Members of the National Panel

As described in Chapter II, in addition to the systematic review of measures, several members of the National Panel prepared reviews or position papers around particular dimensions of program effects. Some of the conclusions regarding the measures they reviewed are quoted below. The comments touch on the range of criteria and considerations discussed above.

Regarding measures of children's social development:

> I have reviewed all of the major studies of Head Start and related programs and all of the instruments used in this research to assess children's social development. I cannot recommend any of these instruments for adoption in this project since in no case did I find satisfactory evidence of the instrument's validity (Carew, 1978, p. 7).

Regarding measures of social-emotional development (including self esteem and reference group orientation):

> For the most part, the only techniques sensitive to socioemotional development in children who range in age from three to seven years, are rating scales and observational methods (Cross, 1978, p. 23).

Regarding Piagetian measures of conservation:

> ... It would seem highly important to establish the extent to which these procedures are applicable (to the Head Start population) (DeAvila, 1978, p. 7).

Regarding measures of fine and gross motor development:

> Indeed, most of the scales of motor development available contain so few items per age level ... that they are unlikely to be discriminating except under circumstances of marked differences between the groups being assessed (Eichorn, 1978, pp. 5-6).

Regarding measures in the Cognitive Domain:

> It is clear that evaluators of Head Start have not taken into account, in their selection of measures, the complex issues underlying the identification of goals and objectives of the program (Laosa, 1978, p. 19).

Regarding the assessment of "social competence" generally:

> It is premature to consider present assessment knowledge as at a point where program evaluation can be applied with certainty (Hilliard & Windsor, 1978, p. 18).

Conclusion

While a lot of progress has been made over the last 15 years in overcoming the practical problems of administering complex batteries of measures to Head Start children in diverse settings across the nation, the major indicators of the quality of child measures have not been up to the standards required for a truly relevant and meaningful evaluation of Head Start program effects. Even the better measures are not completely useful because they appear in a context of measurement batteries that are not adequate when considered as a whole. In searching for improved measures, there has been too little emphasis on the coherence of the battery of measures. In addition to meeting the philosophical and technical criteria, all elements of the battery should be functionally comparable and consistent--coherent logically, conceptually and empirically across all domains and age groups.

New Psychometric Challenges

The basic commitments of this project that most directly influence the technology of measures development include commitments to accommodate local program goals, to accommodate diversity of development among children of different cultural/ethnic groups, to accommodate language differences, to allow for individual differences in patterns of development, and to allow for the expression of behavior in diverse settings. Honoring these commitments means new challenges for determining the reliability and validity of measures. The following discussion summarizes the key ways in which this project will differ from other measures-development efforts in considering instrument validity. This is followed by a discussion of specific applications in relation to typical standards for instrument validity and reliability.

Challenges to Measurement Validity

Accommodating local program emphases and diversity of cultural/ethnic groups. Head Start encourages local projects to plan and develop their programs in accord with local priorities, providing they are consistent with the Performance Standards. There are, and should be, qualitative differences in objectives and emphases among different local programs; and evaluation should be sensitive to such differences. It would be unreasonable, and invalid, to evaluate programs on the basis of objectives they do not seek to attain.

On the other hand, it would be impractical to administer a special set of measures for every local program with objectives and emphases substantially different from those of other Head Start projects. It is practical, however, to devise procedures by which the system of measurement and evaluation takes account of such programmatic differences; and it is important that this be done in aggregating program effects across sites. Although this may be thought of as an evaluation issue rather than a measures development issue, the procedures are discussed here to ensure the availability of appropriate measures to obtain the necessary data for such judgments in evaluations of program effects.

The most feasible approach may be to assign weights to the several dimensions of child development being evaluated at a given site, on the basis of their relative importance in that program, and then to adjust the program's "index of program effect" for each dimension in accord with these weights.

Operationally, procedures such as the following could be followed:

1. Have parents and staff at each site rate each dimension of development being evaluated on a scale of one to ten regarding its relevance to the goals and emphases of the local projects.

2. Calculate the site's rating for each dimension as the mean rating obtained from parents and staff; this then constitutes the "weight" given the dimension by this site.

3. Determine the site's unadjusted "index of program effect" for each dimension.

4. Calculate a weighted aggregation of indices across sites for each dimension, using the weights derived for each site, allowing for differences in sample sizes at the respective sites. The result is an adjusted "index of program effect" for each dimension for the program as a whole (or for whatever group of local projects is being evaluated).

This system of weighting behavioral dimensions on the basis of programmatic relevance offers an approach for making evaluations responsive to programs that vary in objectives and emphases. Such variations may reflect different needs of children, different theories of child development or cultural differences in Head Start communities.

In a predominantly Native-American community, for example, social-emotional goals may focus on cooperation rather than on competition as in some other communities. Similarly, a predominantly Black or Mexican-American community may want its children to know and develop positive attitudes toward its distinctive traditions and values, which, although relevant, may not be stressed in a predominantly white community. Indeed, such cultural/ethnic differences appear in the preliminary analyses of the desirable attainments identified by parents and teachers in the input-workshops.

Accommodating language differences. Over 19 percent of the children in Head Start are of Hispanic background, and 3.6 percent are Native American or Alaskan Natives. Another 1.5 percent of the children in Head Start are from Vietnamese, Chinese, Pacific Islander and other Asian families. All these groups include children who are non-English speaking or who are bilingual.

The emphasis in this project on all measures being sensitive to cultural and other diversity in paths and manifestations of children's development is discussed above. But to the extent that language development itself is being measured, or that language is critical to the behaviors being assessed, measures prepared only in English would be inappropriate for non-English speaking or bilingual children. Also, as ratings of children by parents and staff are recommended as an essential component of the proposed measurement battery, it is important that measures also be prepared in languages in which parents and staff are functional.

Moll (1978) has succinctly described the complexities of assessing language:

Communication involves many extralinguistic factors, such as (the) different social features ..., that may determine its accuracy or success. This suggests

> that a valid assessment approach must allow the children to relate their various communication skills differentially to various tasks, situations and contextual demands. Failure to consider the complexity of the social situation or the context in which language is employed could easily lead to erroneous conclusions regarding children's bilingual communicative competence (Moll, 1978, p. 7).

The National Panel recommended that all measures be developed simultaneously for administration in both English and Spanish. A lag in the development of the Spanish version would necessitate omitting a large segment of the Spanish-speaking Head Start population from the initial validation sample. Also subsequent development of a Spanish version to "match" an already existing English version is a technically inferior and much more costly procedure.

To the extent that funds are available, the Panel has also recommended simultaneous development of measures for other non-English speaking populations that comprise significant numbers of Head Start children.

Accommodating diversity in patterns of development. The selection of a standard of comparison by which to assess children's development is, strictly speaking, an evaluation question; but it impinges directly on the development of measuring techniques, since the data-gathering procedures used in evaluation must yield information relevant to the standard selected. Several issues are involved.

One issue is whether to evaluate program effects on children in terms of *status* attained in the program or extent of *development* in the program, i.e., how far they have progressed in relation to the point at which they entered the program. This is not a technical question. It is a philosophical question involving values, and a theoretical question involving conceptions of child development. Moreover, how the issue is resolved necessarily dictates, in part, the nature of the measuring techniques to be used in evaluation.

As previously noted, the prevailing practice in the field is to assess program effectiveness in terms of the change in status or rank of the children in the program with respect to those on which the tests have been standardized or with respect to some non-treatment group. According to the basic philosophy of this project, the use of "norms" or comparison groups is appropriate only to the extent they reflect development of homogeneous groups of children with similar profiles of entry characteristics. The measures developed for this project will be designed to measure change in children's development during the program, comparing their performance status after the program experience with their status at the time they entered the program. This change would then be appraised in terms of empirically determined development of similar children in similar programs. This evaluative standard thus used might be termed a "dynamic norm"; it is a norm of development (not of status); it is derived from experience in the program being evaluated; it is applied only to homogeneous subgroups of children with a different standard (or "norm") of developmental change for each such homogeneous group. As Saundra Aldridge, President of the National Head Start Staff Association, told her colleagues on the National Panel of this project, it is common for Head Start teachers to say: "Don't just take at face value what my kids *are* doing; look at where they were when they came to me" (Minutes of Panelists' Meeting; Mediax, 1978f).

Accommodating this concern requires attention to problems in measuring change. The problem has received considerable attention in the evaluation and measurement literature. Cronbach and Furby (1970) concluded that it is preferable to rephrase one's research question rather than ask questions regarding gain scores. Linn and Slinde (1977) basically concur, but go on to suggest that pretest scores be considered in the same light as other measures (e.g., entry characteristics) obtained at the time of pretest:

> Where appropriate, regression analyses that treat the pretest no differently from other independent variables (or predictors), and the posttest as the dependent variable avoid many of the difficulties that are introduced by gain scores (Linn & Slinde, 1977, p. 148).

In coping with problems of measuring change for this project the future use of the measures must be kept in mind. When used to examine group differences (either between groups within the Head Start program, or between a Head Start group and a non-Head Start comparison) the major problem to be addressed is that pre-existing group differences are likely given the difficulty of implementing randomized designs for many of the Head Start evaluation questions.

Assessing behavior in a variety of settings. In many cases the desired characteristics identified by parents, teachers and staff specify or imply a particular setting, such as the Head Start classroom; in other cases no setting is indicated and it will be necessary to judge whether it is important or necessary to provide for assessment in more than one setting. Both practical and technical issues are involved in accommodating diverse settings. To the extent that procedures practical for large-scale Head Start evaluations can be devised, it will be critical to then address the issue of across-setting validity. Setting thus becomes one more facet in a study of the generalizability of scores. Because of the central importance of the principles of generalizability theory to this measures-development effort, the implications of the approach are discussed in a separate section below.

Application of Standards of Validity and Reliability

Standards for reliability and validity of measures have received considerable attention from professional associations concerned with measurement and evaluation. As early as 1954 the American Psychological Association published "technical recommendations" for tests. The most recent comprehensive publication, the Standards for Educational and Psychological Tests developed by a joint committee of the American Psychological Association, the American Educational Research Association, and the National Council on Measurement in Education, clearly presents the most basic criteria that all measures should meet.

The application of these standards will not be straight-forward for many of the new measures, however, since these measures will differ in a number of ways from the types of measure on which most approaches to psychometrics have been developed. In addition, the Standards admittedly deal inadequately with criteria specific to measures intended for program evaluation and to the interpretation of scores from content-referenced tests, which some of these new measures may resemble. Nevertheless, developers of new Head Start measures need to appropriately address these commonly accepted standards for tests and measures.

In developing the Standards, the joint committee intended that they "apply to any assessment procedure, assessment device, or assessment aid; that is, to any systematic basis for making inferences about characteristics of people" (APA, 1974, p. 2). Nevertheless, it was recognized by the authors of the Standards that their applicability might vary: "The degree of applicability of individual standards to nontest assessments will vary; developers and users of such assessment procedures should at least observe the spirit of the standards" (APA, 1974, p. 4, emphasis added).

The standards which have the most direct bearing on measures-development activities to Head Start are found in the section on "Standards for Reports of Research on Reliability and Validity" (APA, 1974, pp. 25-55). Standards E1 through E13 deal with validity; Standards F1 through F6 deal with reliability and measurement error. They are reproduced in Table XV.

Validity standards. One of the most comprehensive discussions of issues in measurement validity can be found in Cronbach (1971). In that chapter he points out that "One validates, not a test, but an interpretation of data arising from a specified procedure" (Cronbach, 1971, p. 447). This also represents the first standard (E1) listed in Table XV. Although there are a number of distinguishable uses of the Head Start measures (see the seven uses of the measures listed in Chapter I, p. 19), they share the basic purpose of assessing the effectiveness of the Head Start program. Thus, the various interpretations of data arising from the procedures developed for Head Start in Phase II of this project will have a great deal in common. In fact, one of the reasons this project was launched is that Head Start evaluations have for too long been forced to rely on measures that have been validated for other purposes, such as individual diagnosis or screening. In considering how validity standards for tests and measures apply to the Head Start measures, therefore, we can focus on issues in validating measures for use in Head Start program evaluations. Three types of validity are described in the APA/AERA/NCME Standards: Criterion-related validity, content validity, and construct validity.

Criterion-related validity has a special meaning in the context of Head Start measures development. First, one criterion typically referred to in standard tests of aptitude and achievement--a future measure of ability or achievement--is not a concern of this project; thus, when the Standards refer to predictive validity (e.g., E8) the spirit of that recommendation should be incorporated into measures-development activities, but not the specific concern with using data from these measures to predict some future performance.

The ultimate "criterion" for these measures will, of course, be their ability to correctly identify successful Head Start programs. This cannot be investigated meaningfully until Phase III, but in the Phase II activities, proxy criteria can be examined on a smaller scale. The criteria employed for validating these measures will most likely be other measures of the child's characteristics. It is likely that other measures will be needed because of the restrictions placed on the final versions of the Head Start measures. Requirements, such as being suitable for administration by paraprofessionals, as well as time limitations for the entire battery, place some restrictions on what the final measures will look like. In the process of developing each measure, however, it will be possible to demonstrate its validity in relation to children's performance in a variety of concurrent situations, which might include more extensive observation and professional judgment of the characteristics being measured. In this sense, criterion-related validity assumes great importance for this project, and the Standards concerned with the validity of the criterion measure (E4), the conditions under which information is collected (E6), and data collection procedures (E7) are particularly relevant.

TABLE XV

Reliability and Validity Standards for Educational and Psychological Tests

Validity — General Principles

E1. A manual or research report should present the evidence of validity for each type of inference for which use of the test is recommended. If validity for some suggested interpretation has not been investigated, that fact should be made clear.

E2. A test user is responsible for marshalling the evidence in support of his claims of validity and reliability. The use of test scores in decision rules should be supported by evidence.

Criterion-Related Validity

E3. All measures of criteria should be described completely and accurately. The manual or research report should comment on the adequacy of a criterion. Whenever feasible, it should draw attention to significant aspects of performance that the criterion measure does not reflect and to irrelevant factors likely to affect it.

E4. A criterion measure should itself be studied for evidence of validity and that evidence should be presented in the manual or report.

E5. The manual or research report should provide information on the appropriateness of or limits to the generalizability of validity information.

E6. The sample employed in a validity study and the conditions under which testing is done should be consistent with recommended test use and should be described sufficiently for the reader to judge its pertinence to his situation.

E7. The collection of data for a validity study should follow procedures consistent with the purpose of the study.

E8. Any statistical analysis of criterion-related validity should be reported in the manual in a form that enables the reader to determine how much confidence is to be placed in judgments or predictions regarding the individual.

E9. A test user should investigate the possibility of bias in tests or in test items. Wherever possible, there should be an investigation of possible differences in criterion-related validity for ethnic, sex, or other subsamples that can be identified when the test is given. The manual or research report should give the results for each subsample separately or report that no differences were found.

E10. When a scoring key, the selection of items, or the weighting of tests is based on one sample, the manual should report validity coefficients based on data obtained from one or more independent cross-validation samples. Validity statements should not be based on the original sample.

E11. To the extent feasible, a test user who intends to continue employing a test over a long period of time should develop procedures for gathering data for continued research.

Content Validity

E12. If test performance is to be interpreted as a representative sample of performance in a universe of situations, the test manual should give a clear definition of the universe represented and describe the procedures followed in the sampling from it.

Construct Validity

E13. If the author proposes to interpret scores on a test as measuring a theoretical variable (ability, trait, or attitude), his proposed interpretation should be fully stated. His theoretical construct should be distinguished from interpretations arising on the basis of other theories.

Reliability and Measurement Error--General Principles

F1. The test manual or research report should present evidence of reliability, including estimates of the standard error of measurement, that permits the reader to judge whether scores are sufficiently dependable for the intended uses of the test. If any of the necessary evidence has not been collected, the absence of such information should be noted.

F2. The procedures and samples used to determine reliability coefficients or standard errors of measurement should be described sufficiently to permit a user to judge the applicability of the data reported to the individuals or groups with which he is concerned.

F3. Reports of reliability studies should ordinarily be expressed in the test manual in terms of variances of error components, standard errors of measurement, or product-moment reliability coefficients. Unfamiliar expressions of data should be clearly described, with references to their development.

Comparability of Forms

F4. If two or more forms of a test are published for use with the same examinees, information on means, variances, and characteristics of items in the forms should be reported in the test manual along with the coefficients of correlation among their scores. If necessary evidence is not provided, the test manual should warn the reader against assuming equivalence of scores.

Internal Consistency

F5. Evidence of internal consistency should be reported for any unspeeded test.

Comparisons Over Time

F6. The test manual should indicate to what extent test scores are stable, that is, how nearly constant the scores are likely to be if a parallel form of a test is administered after time has elapsed. The manual should also describe the effect of any such variation on the usefulness of the test. The time interval to be considered depends on the nature of the test and on what interpretation of the test scores is recommended.

With content validity the concern is with the extent to which the behavior tapped in the measurement process itself is representative of the universe of situations or events the measure intends to represent (see Standard E12). If the measuring procedure essentially encompasses the universe, as in a narrow dimension of development, then content validity is not an issue. But in many domains of development the behaviors of interest are only sampled by the measure. The process of defining the universe (or the domains) has been described in Chapter IV. One purpose of this background is to guide the measures developers toward the appropriate content.

In the view of some, content validity, with its focus on the extent to which the content of measure is related to what is intended to be measured, is the most sensible approach. For example:

> After this quarter of a century, which has seen the development of highly complex statistical methods with respect to predictive validity and highly complex logical arguments with respect to construct validity, good old homespun content validity has gradually grown in stature as the most defensible base for many types of psychological tests (Nunnally, 1975, p. 11).

One typically thinks of content validity in connection with achievement tests since they are the most obvious example of measures which sample items from often very large universes. Certainly the child characteristics classified as knowledges and skills fit this example. But it is also likely that measures of attitudes and habits, executive skills and physiological statuses will sample "content" from the relevant universes.

Construct validity is often sought, but seldom accomplished in psychological and educational measurement. To demonstrate that a measure possesses construct validity requires essentially a program of research that seeks to describe the network of relationships among observables, as hypothesized from assumptions relating to the construct (Cronbach, 1971). The characteristics to be measured for Head Start require construct definition to varying degrees. For some, like "count by 5's, 2's and 10's", the characteristic named is very close to the behavior that might be measured; in cases such as this, the measures developers need not be concerned with issues of construct validity. For other characteristics, such as "write creatively" or "persistence" and for all the dimensions and domains identified for measurement, hypotheses are necessary on the relationship of observable events to each of the characteristics, or constructs, to be measured. The structure of domains and dimensions described in Chapter IV provides a starting place for the measure developers. The actual measures development work will require extensive testing of these relationships in order to move toward construct validity during Phase II.

An interesting example of the process (and problems) of construct validation is provided by Shavelson, Hubner and Stanton (1976) for the construct, "self-concept". They discuss the importance of a sound theoretical basis and of a coherent set of research studies, which by means of logical, correlational and experimental analyses, investigate rival interpretations of empirical evidence related to the construct definition.

In addition to all the technical considerations of content, criterion-related and construct validity, the measures for this project need to have what Goodwin has described as "consumer validity" (Goodwin 1978). He notes that the strongest measures in terms of the usual criteria for psychometric validity will lack utility in future Head Start evaluations if key consumer groups such as parents and members of

Congress do not perceive the measures as relevant to their concerns.

Goodwin likens consumer validity to face validity. It is not validity in any technical sense, but rather a value judgment as to whether the measure "looks valid". Having high face validity is no guarantee that the measure is tapping what it "looks like" it's tapping, but it is a desirable feature of measures that are otherwise deemed to show adequate validity.

One function of the Input Workshops held in the early stages of this project is to enhance consumer validity among important Head Start consumers--parents, teachers, staff and public school teachers who've had experience with Head Start children. This has been accomplished for the definition of characteristics to be measured. In Phase II it will be important to continue to receive input of this nature so that the final measures will also appear valid to the ultimate users.

Reliability standards. The basic issue to which most discussions of reliability are aimed is that of the dependability or stability of the measurement. Stanley (1971) provides a systematic treatment of reliability from the perspective of classical test theory. He points out that the basic issue in determining reliability from the classical perspective is to define which factors contribute to "true" variance among individuals and which contribute to "error" variance. Stanley's Table 13.1 (reproduced on the following page as Table XVI) is useful for illustrating some of the potential sources of variance Head Start measures developers will need to be concerned with. Although this list was devised with tests in mind, analagous items can be envisioned for these six sources of variation applied to rating scales, observation techniques, and other types of measurement.

The classical partitioning of score variance into "true" and "error" variance has for some time created discomfort among test developers and measurement specialists. Haney's (1974) discussion of the history of the term "reliability" notes that as early as 1913 Thorndike was offering the view that there is no "error", only varying degrees of probable approximation. When the APA/AERA/NCME Standards were published in 1974 it was acknowledged that:

> The estimation of clearly labeled components of scores variance is the most informative outcome of a reliability study, both for the test developer wishing to improve the reliability of his instrument and for the user desiring to interpret test scores with maximum understanding (APA, 1974, p. 49)

The most systematically developed way of treating reliability consistent with this view is "generalizability theory", a conception of score dependability developed by Cronbach and his associates (Cronbach, Gleser, Nanda, & Rajaratham, 1972). The potential for applying the concepts of generalizability theory in this project is discussed in the following section. First, however, a word about the place of internal consistency reliability in Head Start measures development is in order.

It has become increasingly popular for test developers and users to report internal consistency reliability coefficients as an indicator of the test's "reliability". Such coefficients are only appropriate, however, when the measure is composed of a highly homogeneous set of items. Many measures developed for use in Head Start, particularly observational procedures, will not lend themselves to this type of analysis. When measures are composed of homogeneous items (such as a rating scale that employs several highly similar items relating to the characteristic being rated)

TABLE XVI[a]

Possible Sources of Variance of Scores on a Particular Test

I. Lasting and general characteristics of the individual

- A. Level of ability on one or more general traits, which operate in a number of tests
- B. General skills and techniques of taking tests ("test wiseness" or "test naivete")
- C. General ability to comprehend instructions

II. Lasting but specific characteristics of the individual

- A. Specific to the test as a whole (and to parallel forms of it)
  - 1. Individual level of ability on traits required in this test but not in others
  - 2. Knowledges and skills specific to particular form of test items
  - 3. Stable response sets (e.g. to mark A options more frequently than other options of multiple-choice items, to mark true-false items "true" when undecided, or to choose socially desirable options)
- B. Specific to particular test items
  - 1. The "chance" element determining whether the individual does or does not know a particular fact (sampling variance in a finite number of items, not the probability of his guessing the answer)
  - 2. Item types, such as the data-sufficiency items of the Scholastic Aptitude Test, with which various examinees are unequally familiar (cf. II. A.2)

III. Temporary but general characteristics of the individual (Factors affecting performance on many or all tests at a particular time)

- A. Health
- B. Fatigue
- C. Motivation
- D. Emotional strain
- E. Test-wiseness (partly lasting; cf. I. B)
- F. Understanding of mechanics of testing
- G. External conditions of heat, light, ventilation, etc.

IV. Temporary and specific characteristics of the individual

- A. Specific to a test as a whole
  - 1. Comprehension of the specific test task (insofar as this is distinct from I.B)
  - 2. Specific tricks or techniques of dealing with the particular test materials (insofar as distinct from II.A.2)
  - 3. Level of practice on the specific skills involved (especially in psychomotor tests)
  - 4. Momentary "set" for a particular test
- B. Specific to particular test items
  - 1. Fluctuations and idiosyncrasies of human memory
  - 2. Unpredictable fluctuations in attention or accuracy, superimposed upon the general level of performance characteristic of the individual

V. Systematic or chance factors affecting the administration of the test or the appraisal of test performance
   A. Conditions of testing - adherence to time limits, freedom from distractions, clarity of instructions, etc.
   B. Interaction of personality, sex, or race of examiner with that of examinee to facilitate or inhibit performance
   C. Unreliability or bias in grading or rating performance

VI. Variance not otherwise accounted for (chance)
   A. Luck in selection of answers by sheer guessing
   B. Momentary distraction

[a] Reprinted from Stanley (1971, p. 364, Table 13.1)

internal consistency indices may be useful, but it should not be assumed that they substitute for indices of the measure's reproducibility over time.

The application of generalizability theory. Although the seminal ideas of generalizability theory have been around for at least 40 years (Haney, 1974), the first comprehensive treatment appeared with the publication of The Dependability of Behavioral Measurements by Cronbach et al. in 1972. This approach abandons the classical notions of "true" and "error" scores, and instead analyzes test scores in terms of multiple components of variance. There are many reasons why test scores and scores from other types of measures vary; that is, there are many sources of score variation. These include variation over time, over different forms of a test, over different test administrators, over different raters for a rating scale, over different observers, over scorers, over settings, over certain conditions of testing or rating or observing, and so forth. Each of these sources of variation is termed a "facet" by Cronbach et al.

Four arguments for considering the generalizability approach are listed by Cronach et al.:

1. Explicit consideration of the several facets of a measuring operation dispels ambiguities that were present in, and concealed by, the classical model.

2. The multifacet study can appraise interactions inaccessible to the older methods, and so can improve one's understanding of the measure.

3. One multifacet study answers questions that formerly required several separate sets of data.

4. Multifacet information enables one to design more efficient procedures for collecting data, either for the measurement of individuals or for the determination of group means. (Cronbach, et al., 1972, p. 2).

One of the major issues for measures developers in the Head Start project is to decide which sources of variation need to be accounted for in the particular dimensions being assessed, and which may be either ignored or kept constant. Affholter and Bache (1979) describe the process and rationale for selecting fixed facets and for deciding which facets to vary randomly in the National Day Care Study. Anticipating the future use of measures in Head Start evaluations allows us to make several projections. First, it seems likely that future Head Start evaluations will not exercise control over a number of facets that may affect children's performance on measures in some of the dimensions. These include (a) certain characteristics of the administrators, (b) the sequence in which measures are administered, (c) certain features of the physical surrounding in which tests are administered or in which observations are conducted, (d) time of day of data collection, and so forth.

Second, there are a number of facets that may be controlled or kept relatively constant in future evaluations. These include such facets as (a) the training provided to the administrators of the measures, (b) time of year in which measures are administered, and (c) scoring procedures.

Finally, as spelled out in other sections of this report, there are some facets that must be systematically represented by sampling several conditions if the measures are to meet the several criteria established for them. For example, (a) the cultural/ethnic-group membership of the children will include at least three different "conditions"; (b) the locale will systematically vary from rural to urban and suburban; (c) children's age will be a facet that is systematically studied; (d) for some dimensions the setting will vary, with behaviors being observed in several different settings.

In keeping with the recommendations of the APA/AERA/NCME Standards quoted above, the Head Start measures-development project can greatly benefit by the application of the concepts of generalizability theory to the complex problems of reliability and validity. Rather than following traditional approaches that yield separate indices such as test-retest reliability, internal consistency, interrater agreement, criterion-related validity, and so on, with generalizability theory it is possible to provide a statement about the sources of variation contributing to the scores derived from each of the measures.

Issues in Data Analysis

Throughout this report, as we have described the many ways in which the Head Start approach to measures development will differ from previous efforts, reference has been made to the need for creative approaches to statistical procedures. The standard psychometric methods, while providing minimal standards for measures developers, do not address all the issues that are important for the Head Start measures. In this section of the chapter we briefly discuss some of the major analytic issues that have a bearing on measures development, either because they involve statistical procedures that will be needed in the development work itself, or because they relate to the ultimate use of the measures in ways that may significantly affect the development effort.

Item Analysis

Item analysis is an integral part of any measures-development effort. Most discussions of item analysis, however, assume that items are defined by the procedures typically used in standardized tests of aptitude and achievement. One reason that new approaches need to be entertained is that the "items" of the measures developed for Head Start may include any element of a measure from categories of an observation procedure to questions in a test-like format.

The statistics applied in item analysis are directly related to the assumptions one makes about the nature of the phenomena being measured. Nunnally has represented the traditional approach, arguing that "... the best general approach (with some exceptions) is to develop an instrument that correlates as highly as possible with a particular factor of intellect or personality. This requires that the items be homogeneous with respect to content .... "(Nunnally, 1975, p. 9). In order to select items that correlate well with the test as a whole one must be concerned with maximal discrimination among the individuals taking the test. Linn argues that if one is not concerned with norm-referenced approaches, these traditional item statistics can be ignored:

> For a domain-referenced test, however, such characteristics as item difficulty and discriminating power are generally considered irrelevant. Indeed, when these item statistics are used to select items, they may be viewed as a menace because they can distort the domain representativeness of the test (Linn, 1979, p. 2).

The past decade has seen a shift in strategies employed for item analysis (Baker, 1977). In addition to traditional psychometric techniques based on correlational approaches, there is now an increasing use of statistical estimation procedures, particularly maximum likelihood procedures for estimating item parameters. Developments in this area, which generally is referred to as latent trait theory, item response theory or item characteristic curve theory, attempt to improve on a number of shortcomings of standard psychometric theory (Hambleton, Swaminathan, Cook, Eignor & Gifford, 1978). In particular, Hambleton, et al. cite three problems with standard approaches that latent trait theory strives to overcome:

1. "... values of standard item parameters (item difficulty and item discrimination) are not invariant across groups of examinees that differ in ability. This means that standard item statistics are only useful in test construction for examinee populations very similar to the sample of examinees in which the item statistics were obtained."

2. "... comparisons of examinees on an ability measured by a set of test items comprising a test are limited to situations where examinees are administered the same (or parallel) test items."

3. "Finally, standard testing technology provides no basis for determining what a particular examinee might do when confronted with a test item. Such information is necessary, for example, if a test designer desires ... to design tests with particular characteristics for certain populations of examinees" (Hambleton, et al., 1978, pp. 467-468).

To the extent that latent trait theory provides solutions to these problems, its application should be considered for this project. As a special case of latent trait theory, the Rasch Model has been developed (e.g., Wright & Stone, 1979) and has seen increasing application in educational testing (see Rentz & Rentz, 1979, for a review of recent applications of the Rasch Model). Rentz and Rentz view the Rasch Model as primarily a test construction model, making it particularly relevant to the task of item analysis.

Unfortunately, latent trait theory in general, and the Rasch Model in particular, are not without their critics. Some of the issues involved in the use of these approaches which suggest caution in their application include the extent to which the assumption of independence of items can be met, the sample sizes needed for calculating the item statistics, and the assumption of unidimensionality of the set of items that comprise the measure (Hambleton, et al., 1978). Furthermore, although in principle these approaches may apply to the new generation of measures developed for Head Start, to date their application has been with very traditional standard achievement and ability tests. Thus, the suitability of latent trait approaches should be explored more fully with respect to the particular domains and dimensions of concern to this project.

Carver's discussion of "edumetric tests" should be noted here. Edumetric, as opposed to psychometric, tests are concerned with within-individual growth, not between-individual differences (Carver, 1974). If his analysis can be extended from the level of individual growth to gains made by groups, it may be highly relevant to Phase II of this project.

Profile Analysis

Since all the individual measures of child characteristics that will be developed for use in Head Start are systematically grouped by domains, dimensions, and by process-content clusters, there is interest in making future evaluations capable of depicting the relative performance of programs in the various areas by means of some sort of profile. Rather than reporting only aggregations of scores by dimension or process-content cluster, reporting child development gains in the form of profiles may be useful for depicting qualitative relationships of the various program effects. These would be, as the name of this project suggests, "Head Start Profiles of Program Effects on Children".

The term "profile" is used in this project simply to designate a set of data representing quantitatively the extent to which children in programs develop various competencies, usually presented in some graphic form. The component indices of specific competencies are based on statistical procedures concerned with both the contribution of individual characteristics and comparability between indices.

The use of such profiles of development has proved to be highly illuminating in clinical analysis, but their use in psychometric analysis has not yet been fully developed. Some (e.g., Nunnally, 1975) would argue that efforts at pattern analysis are futile because of the instability of the individual patterns. And Cooley (1971) points out that one of the basic problems to be overcome in existing techniques for analysis of profiles is that the differences for an individual between highly correlated measures will have low reliability. In this project, however, the concern is with group profiles rather than with patterns of individual performance. This may alleviate some of the statistical problems with profile analysis (but, see the discussion of units of analysis below for related concerns).

One of the most important values of profiles for this project is for communicating the extent and nature of program effectiveness in fostering children's development in the several domains and dimensions. Graphic profiles, in a variety of forms, facilitate the interpretation of relationships that would not be so readily apparent from statistical tables showing the same data. For example, a profile graph can show at a glance the relative extent of children's development in the overall program, as well as by domains and by dimensions, making it obvious that the program is having greater impact on development in some dimensions than in others. More sophisticated analyses could even produce profiles that depict the relative extent of development among dimensions for children with different critical entry characteristics; a graph could readily show qualitative differences in the effectiveness of services provided children with various patterns of entry characteristics.

Various formats of the presentation of indices of program effects, and their associated statistical procedures, should be evaluated during Phase II of this project. Figures 29 and 30 are preliminary alternatives for the presentation of overall program effects on all children in the program. Figures 31 and 32 are alternative presentations for reporting effects on groups of children with different characteristics in a single dimension.

Figure 29

Profile of Overall Program Effects

Sample Column Diagram

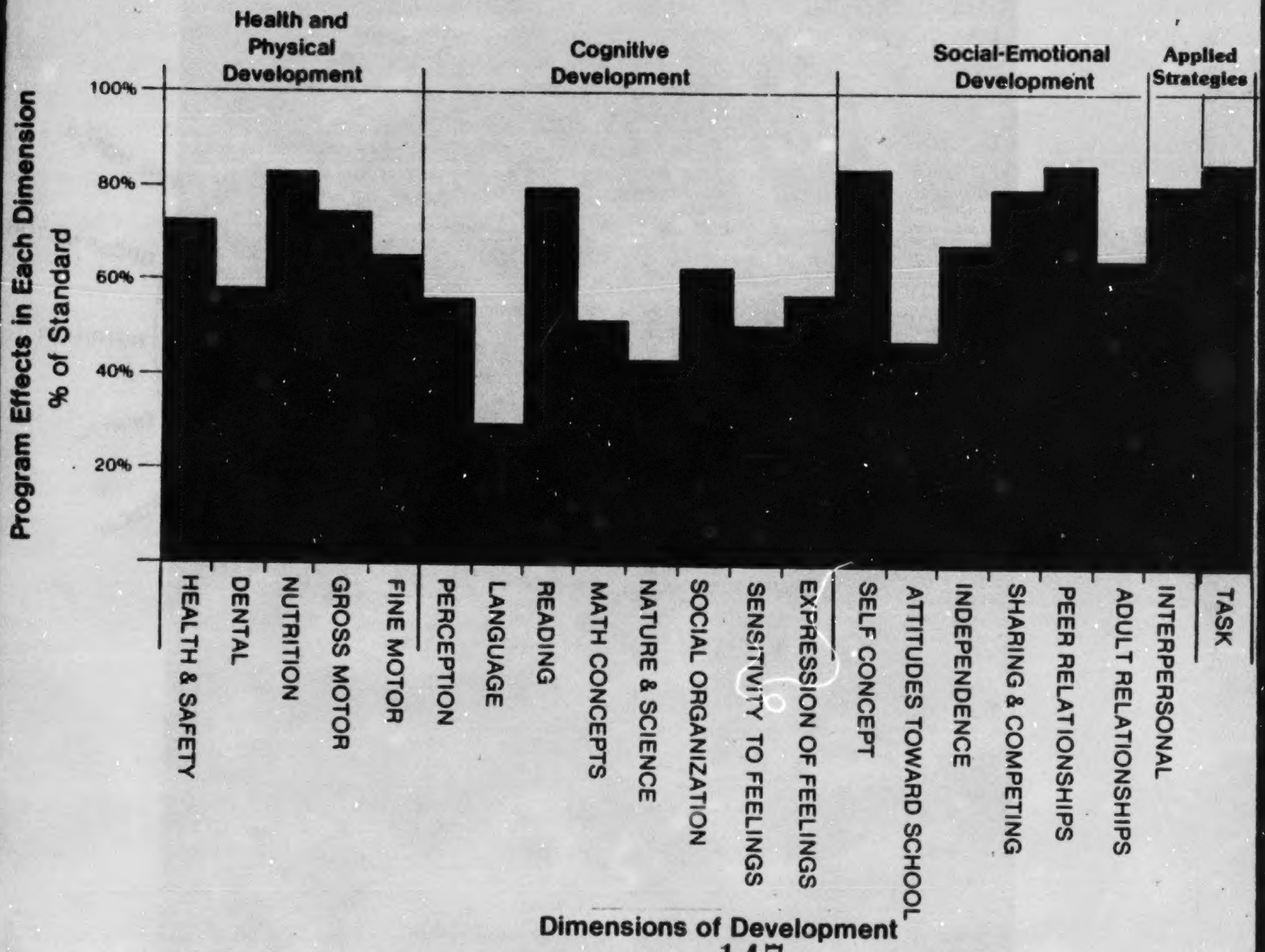


Figure 30

Profile of Overall Program Effects   Sample Sector Graph

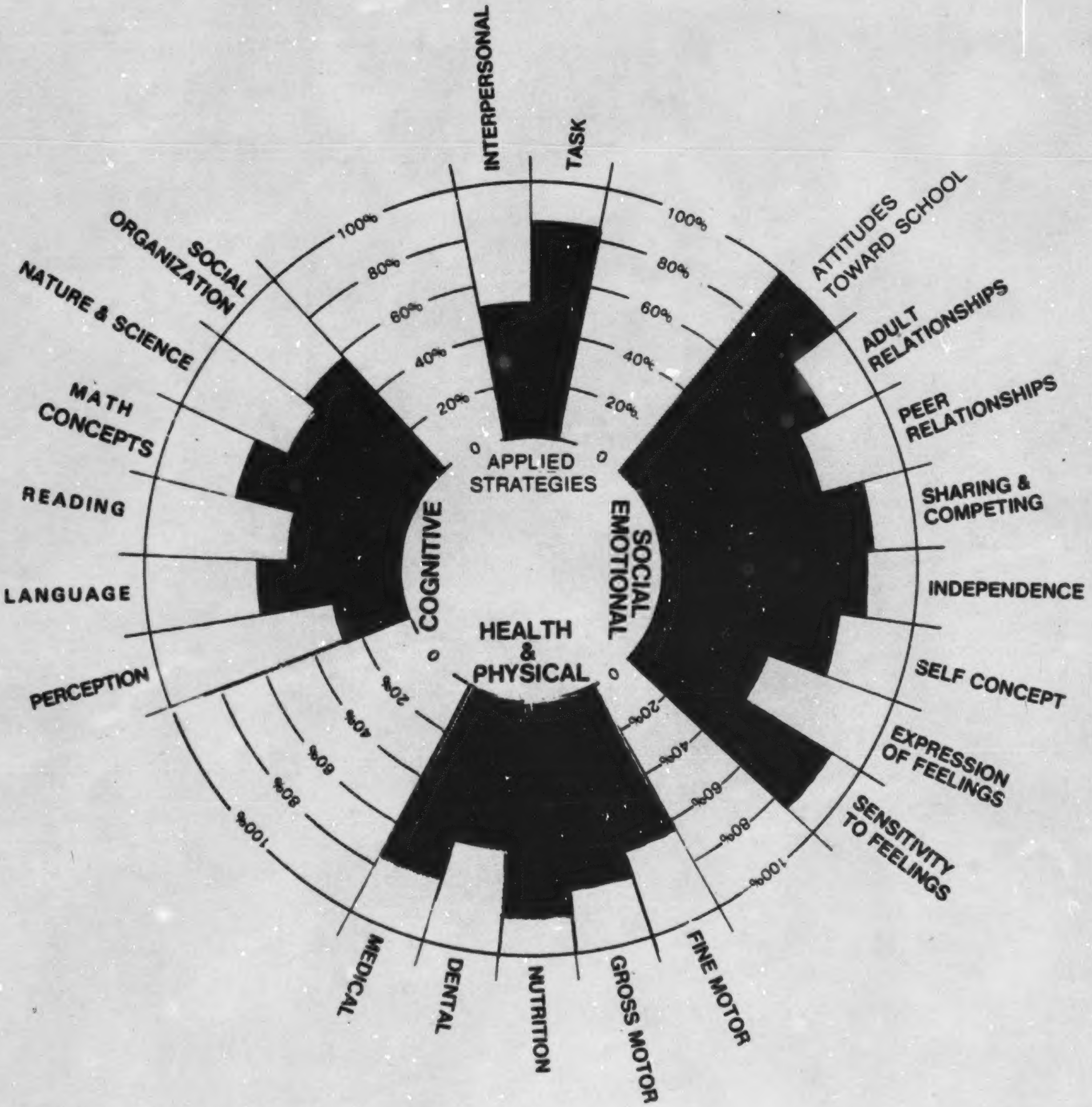


Figure 31

Profile of Program Effects in a Single Dimension
Sample Column Diagram

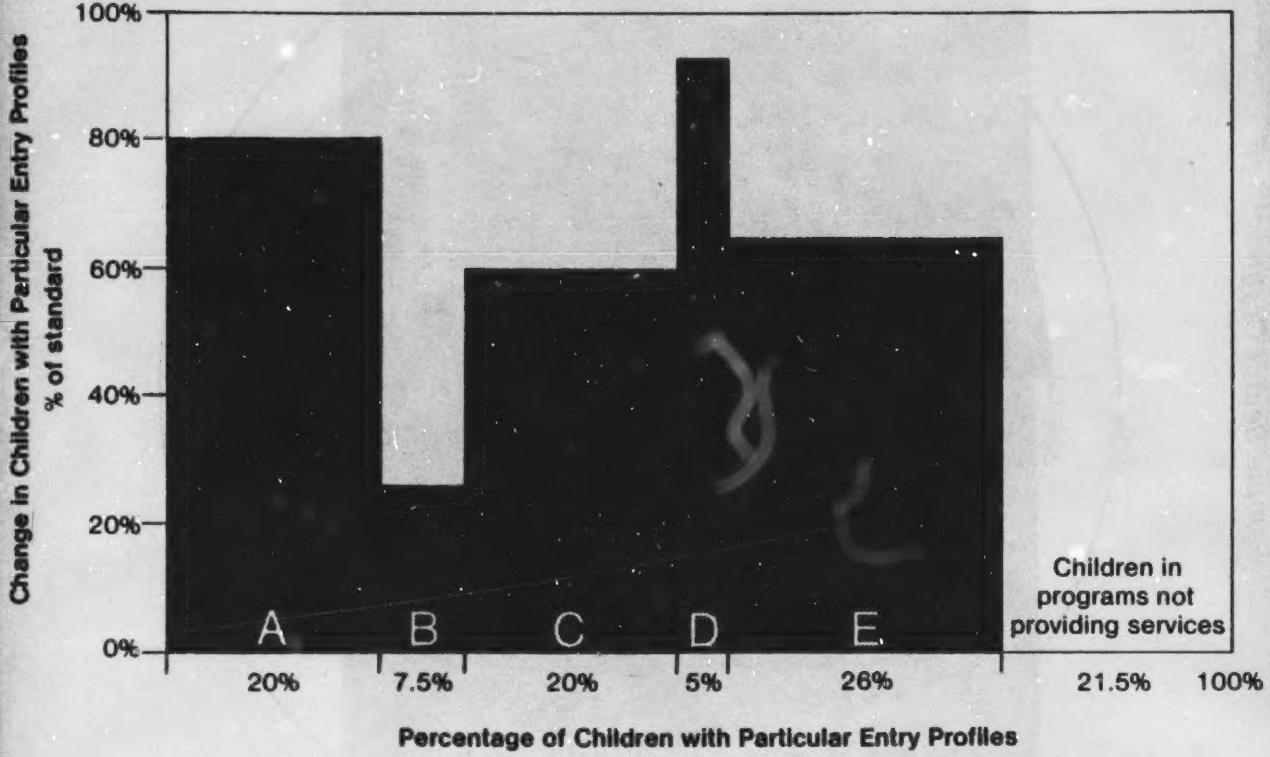
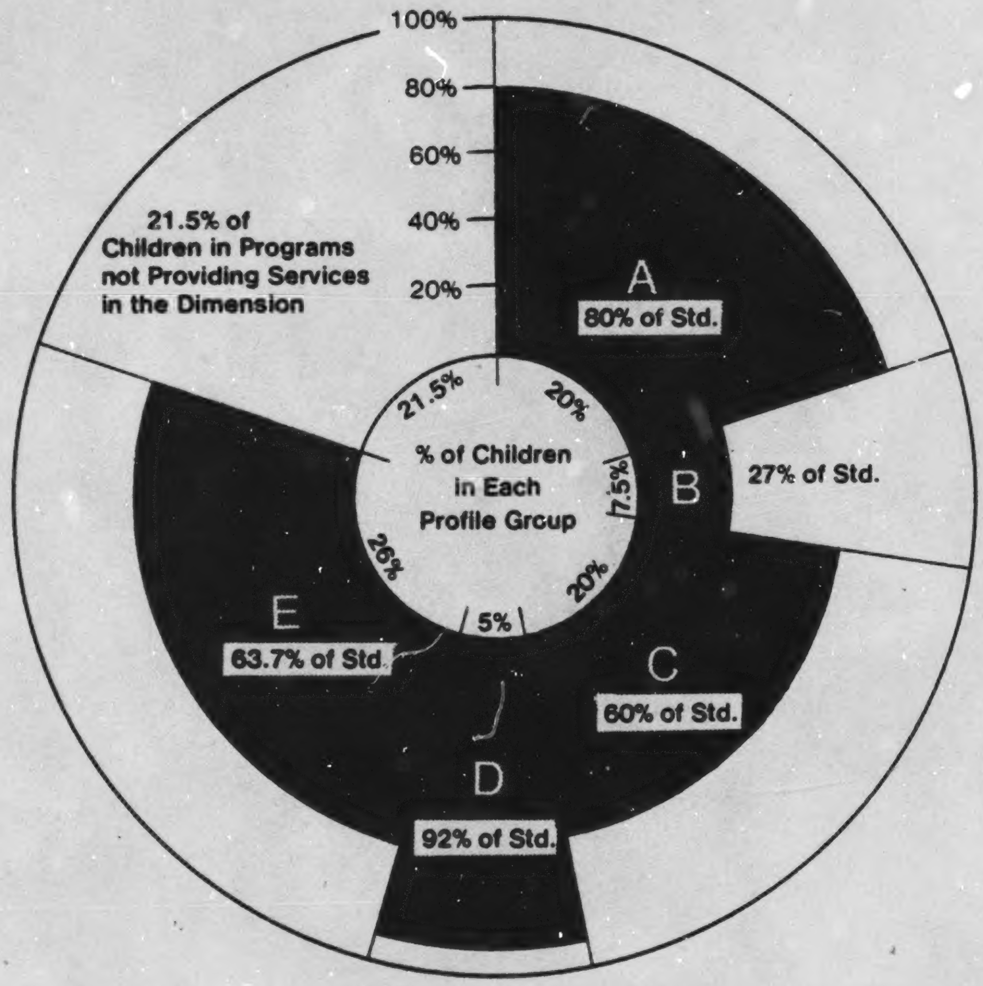


Figure 32

Profile of Program Effects in a Single Dimension
Sample Sector Graph



Reporting Program Effects Without Reference to Norms

As has been discussed at several points in this report, Head Start has for too long been subject to evaluations based on tests whose claim to psychometric adequacy rests largely upon norms which are unrepresentative of the Head Start population. Criterion-referenced measures avoid the use of norms, but may not be totally adequate for Head Start because the diversity of local program goals may make it difficult to ascertain criteria for all programs. If the progress of children could first be determined, without regard to norms and without regard to "control" groups (since control-group designs may not be feasible for some future evaluations), then evaluators and program staff could make judgments regarding the value or importance of the gains made in particular instances.

The "value-added analysis" developed by Bryk and Weisberg (1976) shows promise for applicability to some of the issues of concern to the measurement project. One of the major advantages of their approach is that the model bases its estimations of children's growth on information about the children themselves, not on the growth rates of control-group children, and not on some arbitrary norm. In a recent paper (Bryk, Strenio & Weisberg 1980) the approach has been extended to allow the incorporation of information on background variables, in addition to age, for contributing to the individual growth estimates. Although the ultimate applications of the Head Start measures will require more complex treatments, value-added analysis shows promise of being able to address issues of (a) accommodating individual differences in entry characteristics and cultural/ethnic group membership and (b) avoiding the use of normative comparisons. Furthermore, one limitation of previous applications of the value-added analysis is partially overcome by the fact that pilot-testing of the Head Start measures involves children spanning a wide age range.

Measurement Bias and Fairness

Another major issue with implications for analysis of data from Head Start measures is the importance of developing measures that do not yield scores that are biased by cultural/ethnic group membership of the children. A distinction between fairness and bias has emerged in the testing literature. As described by Merz,

> Fairness is largely an issue arising after a test instrument has been constructed and is being used. Assessing fairness involves determining whether a test used in selection favors choosing members of one group over members of another group. Bias, on the other hand, is an issue relevant to the test construction phase. Items on which a specified group perform differently from others can be identified and either edited or eliminated (Merz, 1977, p. 130).

As described elsewhere in this report, then, the measures-development phase is the appropriate time to address the issue of measurement bias; issues of fairness, to the extent they are concerned with such practices as selecting individuals for jobs or educational opportunities, are not directly at issue here. However, it is likely that attention to item bias at the development phase may in turn alleviate problems of fairness of use in the future.

In the initial development of "items" for the measures, expert judgment plays a role in determining whether the performance of particular groups of children might be affected by the wording, format, content, or concepts involved in the item. Once items are selected and measures are tried out on samples of children, a variety of analytic procedures can be used to assess potential bias. Merz identified and reviewed five techniques used in the test development literature. Each technique involves a somewhat different definition of bias (in terms of the statistical procedures used), and each is likely to yield a somewhat different conclusion regarding the extent of bias present in the items of a particular measure (direct empirical comparisons of these five procedures on the same measures apparently have not been carried out). Merz's logical analysis provides an excellent starting place for Head Start measures developers, however, and suggests approaches that may be most useful for particular types of measures.

Levels of Aggregation

The measures developed in this project will provide indices for each of the hundreds of child characteristics that make up the 21 dimensions being assessed. Obviously, this mass of data must be aggregated in some form to facilitate comprehension by users of evaluation findings. Such aggregation is also necessary to reflect the relative importance of the various characteristics contributing to different areas and levels of child development.

One possibility is to aggregate specific measures of child development characteristics by dimension, providing one index of development for each of the 21 dimensions. Regarding the Domain of Health and Physical Development, for example, a separate numerical index of development would be reported for each of the dimensions--health and safety, dental, nutrition, fine motor, and gross motor. Such indices would also be provided for each of the 16 dimensions that constitute the other three Domains.

As previously noted, an alternative approach to aggregation by dimensions would be to construct a profile showing the relative extent of development in each of the specific characteristics comprising a dimension. Such profiles of development in sub-dimensional characteristics would contain more information than aggregated, quantitative indices. However, given the very large number of profiles this would require, the consumer would find it almost impossible to separate the signal from the noise. Thus, it is essential, first, to calculate an aggregated, quantitative index of development for each dimension, recognizing that it is a composite of many specific characteristics, possibly with varying weights. Once this is done, it will be necessary for full understanding of why an index is "high" or "low" to investigate the configuration of specific characteristics that make up the dimensional index.

On the basis of quantitative indices of program effect for each dimension, composite indices of children's development in each Domain (i.e., Health and Physical, Social-Emotional, Cognitive, and Applied Strategies) can be derived. This would represent a still higher level of aggregation.

Indices of development for Domains could be aggregated further, into an overall measure of program effect. This highest level of aggregation would yield what might be called an "Index of Social Competence Development", since the child characteristics being measured are presumed to define this overall goal of Head Start. Although desired by some policymakers, such an overall index would be of little importance to consumers at the local level. In any case, such aggregation across Domains is logically inappropriate and not recommended.

Although it is possible to aggregate scores to higher and higher levels of generalization, the interpretation of such composites should be guided by at least two caveats. First, each successive level of aggregation provides information farther removed from the primary data, the extent of children's attainment of specific characteristics. This is true even at the dimensional level, and it is increasingly so as aggregation proceeds to the Domain and overall "social competency" levels. Each level of aggregation represents an increasingly abstract construct of development, and less and less the reality of child development.

Second, the meaning of any aggregated index needs to be validated empirically. The characteristics presumed to contribute to each dimension of competency have been derived from the particular behaviors associated in the minds of Head Start parents and staff with desired development in that dimension. These have some face validity in that the characteristics are in effect a definition of the respective dimensions. Nevertheless, it is important that the association of particular characteristics with particular dimensions, and their respective weights, be established empirically through the data obtained by the measures to be developed in this project.

It should be clear that the aggregation problem for the measures contemplated for this project is compounded by the fact that all but one Domain comprises two or more different types of characteristics (see Figure 3, p. 56). Thus for example, the Social-Emotional Domain includes child characteristics classified as knowledges and skills and those classified as attitudes and habits. It is not typical in educational measurement to combine such different types of characteristics into a single index and it may not seem reasonable to do so in this project. But if in real life these different types of characteristics do relate in influencing children's behavior, then it may be possible to achieve more valid measures of constructs by including data on the different types in calculating an index for the construct.

Units of Analysis

Since the proposed measures will be used in Head Start evaluations to assess program effects on groups of children, rather than to assess individual children, it has been stipulated that developers need only concern themselves with the psychometric properties of group means and not of individual scores. Although this is true as a general guideline, the issue warrants further discussion since the classical conception of reliability applied to group means is ambiguous to a certain degree.

Haney (1974) has illustrated the ambiguity that arises when applying classical reliability theory to group scores. Two models of group mean score reliability imply different conceptions of "group". Shaycroft treats group as an aggregation of individuals, so in her reliability formulas the group score is considered an average of all the individuals scores of group members. Wittrock & Wiley, on the other hand, treat the group not as an aggregation but as a unit. As Haney explains it, the approaches differ in the realms to which they generalize:

> Shaycroft's approach, treating the group as an aggregation, in effect generalizes to the group including its members. Wiley's approach, treating the group as a unit, implicitly attempts to generalize to the group itself without including those specific individuals who happen to be members of the group at a particular time (Haney, 1974, p. 44).

The discrepancy between these two approaches arises out of the controversy in classical reliability theory as to which components of variance should be assigned to true score variance and which to error score variance. The procedures of generalizability theory provide an alternative approach to this problem.

Kane and Brennan (1977) point out that classical reliability theory is hampered by its univariate conception of error, whereas generalizability theory depends upon a multivariate conception of error. In their study comparing generalizability coefficients with reliability coefficients calculated for group means (classes) they found that none of three reliability coefficients was equivalent to the generalizability coefficient. Within the framework of generalizability theory, the particular coefficient calculated depends on the purposes of the study, i.e., the sources of variation one wishes to account for. Kane and Brennan conclude by quoting the APA/AERA/NCME Standards to the effect that estimating components of score variance is the most useful form of reliability (see Standard F3 in Table XV), and note that:

> This is equally true whether the unit of analysis is a person or an aggregate of persons, such as a class ... a single generalizability study can replace a number of separate reliability studies (Kane & Brennan, 1977, p. 291).

In considering issues related to units of analysis one needs to consider not only the units on which the analyses are carried out but also the level at which interpretations are to be made. The majority of previous Head Start evaluations have reported group means, group differences and drawn conclusions in terms of group effects. At the same time, however, the analytic techniques employed, such as analysis of variance and analysis of covariance, have been based on individual child scores. The analytic techniques to be used in future evaluations need to be carefully considered before measures developers reach the conclusion that the measures do not need to have demonstrated dependability at the individual level.

Scaling Levels of Development

Inherent in all measurement is some kind of "scale", a device for determining degrees of magnitude or quantity. In this project, scaling is required to assess the extent of children's development in each dimension over time, i.e., the difference between entry and exit indices of development. The problem here is to devise a scale with intermediate points that designate different levels of development in a given dimension, expressed in numerical terms. Ideally, this would be an equal-interval scale on which the difference between, say, indices of 3 and 5 would represent the same increment of development in a dimension as the difference between indices of 7 and 9. Ideally, also, intervals on the scale for one dimension would be comparable to those for all other dimensions. These ideals are rarely achieved in scales of human development. However, it should at least be possible to devise a scale on which "higher" numerical indices are consistently associated with "more" development in the desired direction.

Once such a scale of development is established, the actual gains of children with different entry levels in the "consistently most effective programs" may be used to designate equivalent increments of program effectiveness at different intervals on a scale. Such a procedure will, in effect, produce an equal interval scale of program effectiveness based on data obtained from the programs themselves.

The National Panel recognized the statistical problems related to the development of equal interval scales. However, there has been increasing activity in the development of item analysis techniques which are designed to address this exact problem. The inherent validity and other advantages of such techniques to the outcomes of this project influence a recommendation that they be further considered and evaluated as applicable to the measures being developed during Phase II.

It should be pointed out that it is anticipated that the clusters of characteristics correlating with different points on specific scales may vary among groups of children from varying backgrounds--i.e., with different entry characteristics critical to effects on development in a particular dimension. Thus, there will be qualitative differences in the developmental scales for such groups. It is appropriate that this be so. The diversity of equally valid life styles and values that are represented in the Head Start population should be reflected in the system of measurement by which the program is evaluated.

Data Collection and Measurement Procedures

Many methods of measurement have been used in child development research and in evaluations of early childhood programs. Some of these methods may be ruled out as inappropriate for children of particular ages or for children with particular characteristics. However, as all methods have inherent measurement biases which can only be detected through the use of more than one method of obtaining data on similar characteristics, it is recommended that measures developers be required to use multimethod techniques, at least for analyses during the development and pilot testing phase of this project (Phase II). The particular methods recommended for adoption in subsequent field testing (Phase III) will be selected from those developed, based on considerations of psychometric factors, costs, ease of and time for administration, and overall competeness of coverage of the dimensions.

In considering procedures that may be suitable for measurement and analysis of Head Start program effects, three topics are important: the need for multiple measurement techniques, establishing conditions for eliciting behavior, and data collection issues.

Multiple Measurement Techniques

The selection of particular methods is influenced by the conditions under which behaviors are most appropriately observed. In the history of psychological and educational measurement, a variety of techniques have evolved for meeting a variety of purposes. Most of these techniques have, at some time or another, found their way into evaluations of early childhood programs.

The clinical examination is a standard procedure for measuring status and development in the health area. The clinical approach emphasizes "studying the individual as a unique whole. Specific behaviors are observed and specific traits may be inferred, but the goal is that of understanding (and helping) the particular individual" (English & English, 1958). Although the objective of this project is to provide measures for the assessment of program effects, rather than of individual development, one cannot ignore the fact that this system of measures, and the evaluations in which they are used, may influence services provided individual Head Start children. Therefore, the essential validity of the clinical approach in relation to the detemination of services should not be ignored. It is important that the measures be sufficiently comprehensive to ensure that aspects of the "whole" child and implications for services are not neglected in evaluations. Some clinical methods, modified for program-effects measurement, may be quite appropriate for use in this project, especially in relation to health and physical development.

Archival records, where available, may also be particularly useful sources of evaluative data. Examples include records of attendance, illness, and physical examinations. Some programs may also keep anecdotal records describing individual children's behavior in selected types of situations (e.g., "critical incidents"). Systematic records of these and other types might well supplement evaluative data assembled directly by this project, so long as individual rights to privacy are respected.

Inventories and rating scales for use with parents and teachers afford access to important information not available directly from children, and they can also serve to provide data on generalized constructs to validate aggregation techniques. Moreover, the use of such methods is highly cost-effective. Subjectivity inescapably involved in reports by parents and teachers can be reduced by designing instruments that emphasize factual (and possibly verifiable) information, with a minimum of requests for judgments. In any case, these caregivers know their children better than outside observers can; and much of the relevant information they have cannot be obtained from any other source.

Formal pencil-and-paper tests, which predominate among existing measures of cognitive and social-emotional development, have limited usefulness for assessing preschool children. Reading or writing involves characteristics that are themselves part of the developmental process of children in the age range of concern to this project and would therefore confound data on other characteristics in which they were involved. Where the characteristics being measured are the reading and writing skills themselves, it is recommended that they be observed in settings as close as possible to those in which they are naturally manifested so as to reduce confounding. These procedures may be most useful, however, with children at the upper end of the age range with which this project deals for measuring characteristics that legitimately involve the use of paper and pencil.

Whereas all of the above methods appear to be appropriate for inclusion in the battery to some degree, members of the National Panel feel that significant breakthroughs in validity may be realized in this project only through improved observation techniques. Observational techniques are especially important in that they may yield data on what children actually do in the course of the program being evaluated (and, if feasible, at home and in the community). This allows for assessment of diverse patterns of child development and varying program strategies, a fundamental concern of this project. Observational techniques are also less obtrusive than more direct and formal instruments, and (if applied over a sufficiently long time span) permit assessment of the full range of children's developmental characteristics.

Observation systems also have the potential of gathering in one process assessment information about children's attainments in several dimensions and on several types of characteristics. Such systems could yield important insights on relationships between characteristics.

Three particularly innovative and potentially rewarding approaches to improving observational methods, which are not mutually exclusive, have been recommended by members of the panel. One involves the use of video taped time samples of children's natural behavior as the primary data gathering method (Carew, 1978; 1980). Another recommends development of a coding system to record the elements of all relevant contextual and behavioral variables in sampled observations (Lee, 1978). The third recommmends the development of a series of standard play settings in which children may be observed in semi-natural environments designed to elicit a wide variety of behaviors of interest (Bower, 1978). All have the advantages of permitting confirmation of data, being sensitive to diversity, gathering data on multiple variables simultaneously and have the potential of realizing significant savings in time and cost of administration.

The panel has recommended that developers of measures be encouraged to explore the feasibility of these innovative approaches to observational methods in the development and pilot testing undertaken in Phase II of this project.

Conditions for Eliciting Behavior

Two considerations are particularly important in reaching decisions about how to measure children's development and behavior--the degree to which the measurement setting controls the child's behavior and the extent to which it is necessary to obtain measures of behavior that is typical for the child or behavior that is an index of some maximal level of performance. These two considerations represent separate but interrelated decisions in the measures development process.

In considering the control exerted by the measuring setting, one extreme is a highly standardized situation with restricted stimuli and restricted opportunities for the child to respond. Examples are paper-and-pencil tests, tests with multiple-choice response formats and measures of physical status. The other extreme is a completely "natural" setting in which the evaluator exercises no control over the stimulus situation and there are no external restrictions in the child's behavior (though there may be limits in what is recorded by the evaluator). In between these extremes are measurement settings that create varying degrees of control over either the stimuli or the response opportunities.

The measurement setting that is devised for particular dimensions must also be related to the nature of the desired characteristics. Some characteristics require estimating the child's typical performance (e.g., "know consequences of one's actions")--the measurement concern is determining what a child usually does without particular urging or special conditions to heighten the level of response. Other characteristics suggest that it is the "best" performance, or some maximal level of performance that's to be measured (e.g., the desired characteristic, "can tell details of picture observed", indicates that the child should be able to perform in certain ways when called upon, but may not necessarily need to exhibit the characteristic on his or her own in everyday situations). Some characteristics named in the Input Workshops are ambiguous in this regard. For example, "have a good memory for information" may mean either that children should be able to exhibit "good memory" under highly controlled conditions that maximize the chance to demonstrate memory (e.g., by eliminating distractions or controlling particular properties of the stimuli), or that children should use their memory capacity in day-to-day situations; or, both may be implied by the named characteristic.

The developmental characteristics of children identified for assessment in this project call for both types of measures. Where children's maximal performance is to be evaluated, a carefully controlled, standardized setting is required to increase the likelihood that the child's performance is the best he or she is capable of. The familiarity of the setting, its physical characteristics, the characteristics of the person administering the measure, and the time allotted for repsonses, among other factors, can influence children's performance. Where children's typical performance is to be evaluated, however, highly standardized settings are not desirable. In this case the very factors that create the standard setting make it unusual for the child, increasing the likelihood that atypical behavior will be observed. To allow for the assessment of both typical and maximal behavior, it is recommended that measurement be conducted in both natural and standard settings as appropriate for the characteristics being measured.

Data Collection Issues

Although data collection in any future Head Start evaluation will involve a large number of activities that conceivably impinge on the validity and reliability of the measures, two issues can be anticipated because of their central importance--the time required for the child to participate in measurement activities and the nature of the personnel who will conduct the data collection.

Data collection time, when it directly involves Head Start children, staff or parent should be kept to a minimum so as to reduce the burden on these program participants (maintaining reasonable evaluation costs is another problem for large-scale evaluations that relates to time required for measurement). Making efficient use of available measurement time takes on added importance when one considers the large number of child characteristics to be measured.

The guideline adopted for this project is to restrict the amount of time that an individual child can spend in responding to measuring instruments. The specific recommendation is that a maximum of one hour (divided into two half-hour sessions) be permitted for directly involving children during any data collection period (i.e., at pre-test or post-test) at the Head Start level. (A somewhat longer time may be tolerated for children in the first and second grade.) This recommendation will be implemented by identifying, from all the measures developed by the end of Phase II, clusters of measures that can be administered within this timeframe.

This limitation, of course, is not a restriction on the total amount of information that can be obtained regarding a particular child. Other sources, such as ratings by parents and teachers, observations of program activities, and examinations of existing records, can substantially augment information obtained through direct participation by the child. In subsequent program evaluations, the measurement burden can be further lessened through matrix-sampling techniques so there is reason to plan for the development of measures that would total more than one hour of child response time.

For a number of reasons it has been recommended that paraprofessional personnel be used in any testing or observation that is conducted as part of Head Start evaluations. The use of paraprofessisonals has many advantages. Drawn from backgrounds similar to those of the populations in the programs being evaluated, they are likely to "sense" and react appropriately to subtle characteristics of the setting and of children's behavior. When effectively trained in data-gathering techniques, they may have less difficulty than "outsiders" in establishing rapport with parents and teachers from whom certain types of information will be needed. Moreover, their participation in the measurement and evaluation process is likely to enhance its consumer validity.

This requirement does place some limitations on the nature of the measures that can be developed. For example, administration procedures need to be clearly specified and the process of recording responses must not require sophisticated psychodynamic judgments. Previous Head Start evaluations have dealt successfully with this issue, and if this requirement is acknowledged from the outset the chances are even greater for developing reliable and valid measures that can be administered by paraprofessionals.

VI

PERSPECTIVES AND REASONABLE EXPECTATIONS

This measures-development project envisions long-term outcomes that are difficult to realize -- the identification of programs that are most effective for children with diverse needs, goals and paths of development. What is certainly accomplishable, however, is a significant improvement in the "tools" available for evaluating the impact of Head Start on children.

The immediate objective is clear -- a comprehensive and coherent battery of measures that will yield valid, reliable and reasonably precise information about the development of all Head Start children in those areas the program seeks to affect. Thus, this project seeks consciously to avoid many of the pitfalls of measuring techniques heretofore employed in evaluations of Head Start. Notable among them are the use of content irrelevant to the program's objectives, the glossing over of diverse but equivalent patterns of development among cultural groups, the consequent sterotyping of minority-group children on the basis of so-called "national norms", the focus on status of achievement rather than change in development, the use of disparate instruments that preclude overall comparability, and, of course, questionable psychometric properties.

If future Head Start evaluations are to overcome these shortcomings, dramatic changes are necessary in the measures for assessing program impact on children. Many of these changes have important implications for the ways in which educational researchers, evaluators and psychometricians think about the technical qualities of the measures; even more importantly, they can affect the way policymakers, administrators and practitioners think about the effectiveness of programs for children. This report has attempted to identify a feasible approach to accomplishing these very important changes.

Although compromises will continue to be required in realizing all of the long-term outcomes envisioned in this project, one result is reasonably certain. Future evaluations of program effects on young children will have available a battery of measures that is significantly less compromising of the developmental needs of children.

REFERENCES

Affholter, D. P., & Bache, W. L. The application of generalizability theory to observation-based dependent and independent measures in the National Day Care Study. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, April 1979.

American Psychological Association. Standards for educational and psychological tests. Washington, D.C.: American Psychological Association, 1974.

Anderson, S., & Messick, S. Social competency in young children. Princeton: Educational Testing Service, 1973.

Austin, J.E. The perilous journey of nutrition evaluation. American Journal of Clinical Nutrition, 1978, 31, 2314-2338.

Baker, F. B. Advances in item analysis. Review of Educational Research, 1977, 47, 151-178.

Bibby, G. The cariogenicity of snack foods and confections. Journal of American Dental Association, 1975, 90, 121.

Bloom, B. S. (Ed.) Taxonomy of educational objectives: Handbook I: Cognitive domain. New York: David McKay Company, 1956.

Bloom, B.S. Stability and change in human characteristics. New York: Wiley, 1964.

Bower, E. Assessment of personal and social growth factors through play. Westport, Connecticut: Mediax Associates, Inc., 1978.

Bronson, M. B. Manual for the executive and social skill profile for preschool children. Cambridge, Massachusetts: Graduate School of Education, Harvard University, 1973.

Bryk, A. S., & Weisberg, H. I. Value-added analysis: A dynamic approach to the estimation of treatment effects. Journal of Educational Statistics, 1976, 1, 127-155.

Bryk, A. S., Strenio, J. F., & Weisberg, H. I. A method for estimating treatment effects when individuals are growing. Journal of Educational Statistics, 1980, 5 5-34.

Buros, O. K. (Ed.) The sixth mental measurements yearbook. 2 Vols. Highland Park, New Jersey: Gryphon Press, 1965.

Buros, O. K. (Ed.) The seventh mental measurements yearbook. 2 Vols. Highland Park, New Jersey: Gryphon Press, 1972.

Busa, J. Head Start Performance Indicators. Paper presented at the National Head Start Association Meeting, Miami, Florida, May 21-24, 1980.

Butler, A. L., Gotts, E. E., Quisenberry, N., & Thompson, R. P. Literature search and development of an evaluation system in early childhood education. Bloomington, Indiana: Indiana University, 1971.

Carew, J. V. Head Start profiles of program effects on children. Position paper for Mediax ACYF Project. Westport, Connecticut: Mediax Associates, Inc., 1978.

Carew, J. V. Comments of panelist Jean Caraw on the draft Options Paper. March, 1980.

Carver, R. P. Two dimensions of tests: Psychometric and edumetric. American Psycholotist, 1974, 29, 512-518.

Cazden, C. B. Evaluation of learning in preschool education: Early language development. In B. S. Bloom, J. T. Hastings, & G. F. Madaus (Eds.), Handbook on formative and summative evaluation of student learning. New York: McGraw-Hill, 1971.

Cicirelli, V.G., Cooper, W. H., & Granger, R. L. The impact of Head Start: An evaluation of the effects of Head Start on children's cognitive and affective development. Washington, D.C.: Office of Economic Opportunity, 1969.

Ciminero, A. R. Handbook of behavioral assessment. New York: Wiley, 1977.

Collins, R.C. Children and society: Child development and public policy. Unpublished manuscript, 1977.

Collins-Sanford, J. Recommendations of dimensions to be measured in the area of nutritional development of 3 to 7 year olds. Westport, Connecticut: Mediax Associates, Inc., October, 1978.

Cooley, W.W. Techniques for considering multiple measurements. In R. L. Thorndike (Ed.), Educational measurement. (2nd Ed.) Washington, D.C.: American Council on Education, 1971. Pp. 601-622.

Coopersmith, S. The antecedents of self-esteem. San Francisco: Freeman, 1967.

Cronbach, L. J. Test validation. In R. L. Thorndike (Ed.), Educational measurement. (2nd Ed.) Washington, D.C.: American Council on Education, 1971. Pp. 443-507.

Cronbach, L. J., & Furby, L. How we should measure "change" -- or should we? Psychological Bulletin, 1970, 74, 68-80.

Cronbach, L. J., Glaser, G., Nanda, H., & Rajaratnam, N. The dependability of behavioral measurements: Theory of generalizability for scores and profiles. New York: Wiley, 1972.

Cross, W. E., Jr. Assessment of emotional development in children participating in Head Start programs. Westport, Connecticut: Mediax Associates, Inc., 1978.

Datta, L., McHale, C., & Mitchell, S. The effects of the Head Start classroom experience on some aspects of child development: A summary report of national evaluations, 1966-1969. Washington, D.C.: Office of Child Development, 1976.

DeAvila, E. A. *Review of tests available for Piagetian conservation tasks*. Westport, Connecticut: Mediax Associates, Inc., 1979.

Department of Health, Education & Welfare. Request for proposal for development of measures of social competence. Washington, D.C.: Department of Health, Education and Welfare, January 21, 1977.

Edgar, W. M., Bibby, B. G., Mundorff, S., & Rowley, J. Acid production in plaques after eating snacks: modifying factors in foods. *Journal of American Dental Association*, 1975, *90*, 418.

Eichorn, D. H. *Head Start profiles of program effects on children: Motor development*. Westport, Connecticut: Mediax Associates, Inc., 1978.

Ellen, R. P. Microbiological assays of dental caries and periodontal disease susceptibility. *Oral Sciences Review*, 1976, 8, 3-23.

English, H. B., & English, A. *A comprehensive dictionary of psychological and psychoanalytical terms*. New York: David McKay, 1958.

Exner, J. E., Jr., & London, H. (Eds.), *Dimensions of personality*. New York: Wiley, 1978.

Fehr, F. R. V. The effects of fluorides on the caries resistance of enamel. *Acta Odonta Scand.*, 1961, 19, 431-442.

Frankenburg, W. K., & Dodds, J. B. The Denver Developmental Screening Test. *Journal of Pediatrics*, 1967, *71*, 181.

Goodwin, W. L. *The measurement of children's motivation as a program outcome*. Westport, Connecticut: Mediax Associates, Inc., 1978.

Gordon, S. G., in collaboration with C. DeAngelis. *Head Start profiles of program effects on children -- Medical and dental domain*. Westport, Connecticut: Mediax Associates, Inc., 1978.

Hambleton, R. K., & Cook, L. L. Latent trait models and their use in the analysis of educational test data. *Journal of Educational Measurement*, 1977, *14*, 75-96.

Hambleton, R. K., Swaminathan, H., Cook, L. L., Eignor, D. R., & Gifford, J. A. Developments in latent trait theory: Models, technical issues, and applications. *Review of Educational Research*, 1978, *48*, 487-510.

Haney, W. The dependability of group mean scores. Unpublished qualifying paper. Cambridge, Massachusetts: Harvard Graduate School of Education, 1974.

Hilgard, E. G. *Theories of learning*. New York: Appleton-Century-Crofts, 1978.

Hilliard, A. G., III., & Windsor, K. *I'm back!!! Children's social behavior: Assessment considerations and criteria*. Westport, Connecticut: Mediax Associates, Inc., 1978.

Hoepfner, R., Stern, C., & Nummedal, S. G. *CSE-ECRC preschool/kindergarten test evaluations*. Los Angeles: UCLA Graduate School of Education, Center for the Study of Evaluation and Early Education Research Center, 1971.

Hunt, J. McV. *Intelligence and experience*. New York: Ronald Press Co., 1961.

Hunt, J. McV. Letters from J. McVicker Hunt to Herman P. Taub. Unpublished. Westport, Connecticut: Mediax Associates, Inc., February, 1979.

International Training Consultants. *Developmental assessment instruments: Survey and identification of Head Start practices and needs: Final report*. Denver, Colorado: International Training Consultants, March 13, 1978.

Johnson, O. G. *Tests and measurements in child development: Handbook II*, Vols. 1 and 2. San Francisco: Jossey-Bass, 1976.

Johnson, O. G., & Bommarito, J. W. *Tests and measurements in child development: Handbook I*. San Francisco: Jossey-Bass, 1971.

Kamii, C. K. Evaluation of learning in preschool education. In B.S. Bloom, J. T. Hastings, & G. F. Madaus (Eds.), *Handbook on formative and summative evaluation of student learning*. New York: McGraw-Hill, 1971.

Kane, M. T., & Brennan, R. L. The generalizability of class means. *Review of Educational Research*, 1977, *47*, 267-292.

Katz, L., & Jacobson, E. *Measures of development among young children: Socio-emotional domain*. Westport, Connecticut: Mediax Associates, Inc., 1978.

Laosa, L. M. *Evaluating the effects of Head Start on children's cognitive development*. Westport, Connecticut: Mediax Associates, 1978.

Lazar, I., Hubbell, V. R., Murray, H., Rosche, M., & Royce, J. *The persistance of preschool effects: A long-term follow-up of fourteen infant and preschool experiments*. Ithaca, New York: Cornell University, September 1977.

Lee, L. C., assisted by Dopyera, J, and Lay, M. *Assessing communicative competence via the repertoire model*. Westport, Connecticut: Mediax Associates, Inc., 1978.

Linn, R. L. Test design and analysis for the measurement of educational achievement. Paper prepared for the 1979 Invitational Conference on Testing Problems. New York, October 27, 1979.

Linn, R. L., & Slinde, A. The determination of the significance of change between pre- and posttesting periods. *Review of Educational Research*, 1977, *47*, 121-150.

Love, J. M., Granville, A. C., & Smith, A. G. *Final report of the PDC feasibility study, 1974-1977*. Ypsilanti, Michigan: High/Scope Educational Research Foundation, April, 1978.

Love, J. M., Wacker, S., & Meece, J. *A process evaluation of Project Developmental Continuity, interim report II, Part B: Recommendations for measuring program impact*. Ypsilanti, Michigan: High/Scope Educational Research Foundation, June, 1975.

Love, J. M., Nauta, M. J., Coelen, C. G., Hewett, K., & Ruopp, R. R. *National Home Start evaluation: Final report*. Ypsilanti, Michigan: High/Scope Educational Research Foundation and Abt Associates Inc., March, 1976.

Mann, A. J., Harrell, A., & Hurt, M., Jr. Working draft: A review of Head Start research since 1969 and an annotated bibliography. Washington, D.C.: Social Research Group, The George Washington University, December, 1976.

Marthaler, T. M. A Standardized system of recording dental conditions. *Helv. Odont. Acta*, 1966, *10*, 1-18.

Mediax Associates, Inc. *Compendium of information gathering techniques utilized in previous Head Start research projects*. Vols. 1 and 2. Westport, Connecticut: Mediax Associates, Inc., June, 1978. (a)

Mediax Associates, Inc. *Composite list of developmental objectives (coded) for children aged three to seven*. Westport, Connecticut: Mediax Associates, Inc., December, 1978. (b)

Mediax Associates, Inc. *Current issues in the development project for Head Start Profiles of Program Effects on Children*. Westport, Connecticut: Mediax Associates, Inc., August, 1978. (c)

Mediax Associates, Inc. *Descriptions of information gathering techniques used in previous Head Start programs*. Westport, Connecticut: Mediax Associates, Inc., July 1978. (d)

Mediax Associates, Inc. *ERIC Literature indexed to dimensions of child development*, Vol. 3. Westport, Connecticut: Mediax Associates, Inc., 1978. (e)

Mediax Associates, Inc. *Minutes of the panelists' meeting: 1978*. Westport, Connecticut: Mediax Associates, Inc., 1978. (f)

Mediax Associates, Inc. *Original group inputs: Parent-staff-teacher input workshops*. 10 Vols. Westport, Connecticut: Mediax Associates, Inc., 1978. (g)

Mediax Associates, Inc. *Panelist guide*. Westport, Connecticut: Mediax Associates, Inc., January, 1978. (h)

Mediax Associates, Inc. *Preliminary interim report on child development attainments identified by scholars for evaluating program effects*. Westport, Connecticut: Mediax Associates, Inc., April 10, 1978. (i)

Mediax Associates, Inc. *Report of attainments identified in input workshops*. Westport, Connecticut: Mediax Associates, December, 1978. (j)

Mediax Associates, Inc. *Report on process of input workshops*. Westport, Connecticut: Mediax Associates, Inc., August, 1978. (k)

Mediax Associates, Inc. *Resource reference book - test collection: Educational Testing Service*. Volume I. Westport, Connecticut: Mediax Associates, Inc., *previous Head Start programs*. Westport, Connecticut: Mediax Associates, Inc., July 1978. (d)

Mediax Associates, Inc. *ERIC Literature indexed to dimensions of child development*, Vol. 3. Westport, Connecticut: Mediax Associates, Inc., 1978. (e)

Mediax Associates, Inc. *Minutes of the panelists' meeting: 1978*. Westport, Connecticut: Mediax Associates, Inc., 1978. (f)

Mediax Associates, Inc. *Original group inputs: Parent-staff-teacher input workshops*. 10 Vols. Westport, Connecticut: Mediax Associates, Inc., 1978. (g)

Mediax Associates, Inc. *Panelist guide*. Westport, Connecticut: Mediax Associates, Inc., January, 1978. (h)

Mediax Associates, Inc. *Preliminary interim report on child development attainments identified by scholars for evaluating program effects*. Westport, Connecticut: Mediax Associates, Inc., April 10, 1978. (i)

Mediax Associates, Inc. *Report of attainments identified in input workshops*. Westport, Connecticut: Mediax Associates, December, 1978. (j)

Mediax Associates, Inc. *Report on process of input workshops*. Westport, Connecticut: Mediax Associates, Inc., August, 1978. (k)

Mediax Associates, Inc. *Resource reference book - test collection: Educational Testing Service*. Volume I. Westport, Connecticut: Mediax Associates, Inc., June, 1978. (l)

Mediax Associates, Inc. *Selected current literature series: Dimensions of child development*, Vol. 2, Nos. 1 to 32. Westport, Connecticut, Mediax Associates, Inc., 1978. (m)

Mediax Associates, Inc. *A system for selecting characteristics to measure Head Start program effects*. Westport, Connecticut: Mediax Associates, Inc., 1980.

Merz, W. R. Test fairness and test bias: A review of procedures. In M. Wargo & D. Green (Eds.), *Achievement testing of disadvantaged and minority students for educational program evaluation*. New York: McGraw-Hill, 1977. Pp. 129-151.

Minnuchin, P., Biber, B., Shapiro, E., & Zimiles, H. *The psychological impact of schooling experience*. New York: Basic Books, 1969.

Mischel, W. *Personality and assessment*. New York: Wiley, 1968.

Moll, L. C. The importance of the social situation in assessing bilingual communicative performance. *Quarterly Newsletter of the Laboratory of Comparative Human Cognition* (University of California, San Diego), 1978, *1*(1), 5-8.

National Institute for Advanced Studies. *Analysis of 1978-79 Head Start performance indicators*. Washington, D.C.: National Institute for Advanced Studies, 1980.

Nunnally, J. C. Psychometric theory--25 years ago and now. *Educational Researcher*, 1975, *4*, 7-21.

Office of Child Development. Final reports on project Head Start. National evaluation projects, interim listing. Washington, D.C.: U.S. Department of Health, Education and Welfare, 1975. (a)

Office of Child Development. Head Start Program Performance Standards. (OCD Notice N-30-364-4) Washington, D.C.: U.S. Department of Health, Education and Welfare, Office of Child Development, 1975. (b)

Quick, Alton, Little, Thomas, and Campbell. Memphis Comprehensive Developmental Scales. Memphis, Tennessee: Memphis State University, 1972.

Rader, J. R., Thouvenelle, S., Vogel, R. J., & Davis, S. Evaluation of the process of mainstreaming handicapped children into Project Head Start. Phase II draft interim report. Silver Spring, Maryland: Applied Management Science, Inc., March 24, 1978.

Raizen, S., & Bobrow, S. Design for a national study of social competence in Head Start children. Santa Monica, California: Rand Corporation, 1974.

Rentz, R. R., & Rentz, C. C. Does the Rasch Model really work? A discussion for practitioners. Educational Testing Service, ERIC TM Report 67, 1979.

Robinson, J. P., & Shaver, P.R. Measures of social psychological attitudes. Ann Arbor, Michigan: University of Michigan, Institute for Social Research, 1974.

Royster, E. C., & Larson, J. C. Executive summary of a national survey of Head Start graduates and their peers. Cambridge, Massachusetts: Abt Associates, 1978.

Royster, E., Larson, J. C., Ferb, T.E., Fosberg, S., Nauta, M., Nelson, B., Takata, G., & Ferb, K. A national survey of Head Start graduates and their peers. Cambridge, Massachusetts: Abt Associates, 1977.

Ruopp, R. National day care study. Second annual report, 1975-1976, phase II results and phase III design Cambridge, Massachusetts: Abt Associates, Inc., 1977.

Ruopp, R., Travers, J., Glantz, F. & Coelen, C. National Day Care Center Study. Cambridge, Massachusetts: Abt Associates, 1977.

Ruopp, R., Travers, J., Glantz, F. & Coelen, C. Children at the center: Final report of the National Day Care Study. Volume I. Cambridge, Massachusetts: Abt Associates, Inc., March, 1979.

Ryle, A. Measuring the effect of Head Start programs on future academic achievement. Westport, Connecticut: Mediax Associates, Inc., June, 1978.

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. Self-concept: Validation of construct interpretations. Review of Educational Research, 1976, 46, 407-441.

Silva, M. N. Teaching Head Start children survival skills in the mainstream society. Westport, Connecticut: Mediax, Associates, Inc., 1978.

Shaw, M. E., & Wright, J. M. Scales for the measurement of attitudes. New York: McGraw-Hill, 1967.

Shipman, V. C., Boroson, M., Bridgeman, B., Gant, J., & Mikovsky, M. Disadvantaged children and their first school experiences: Notable early characteristics of high and low achieving black low-SES children: ETS-Head Start longitudinal study. Princeton: Educational Testing Service, 1976.

Stanley, J. C. Reliability. In R. L. Thorndike (Ed.), Educational measurement. (2nd Ed.) Washington, D.C.: American Council on Education, 1971. Pp. 356-442.

Stebbins, L., et al. Evaluation design and implementation plan for Head Start health evaluation. Cambridge, Massachusetts: Abt Associates, 1978.

System Development Corporation. Technical Memorandum: Effects of different Head Start program approaches on children of different characteristics: Report on analysis of data from 1966-67 and 1967-68 national evaluations. Santa Monica, California: System Development Corporation, August 29, 1972.

Taub, H. P. Perspectives for the measurement of functional stress range. Westport, Connecticut: Mediax Assoicates, Inc., 1980.

Travers, R. M. Taxonomies of educational objectives and theories of classification. Educational Evaluation and Policy Analysis, 1980, 2, 5-23.

Vogel, J., & Rader, J. R. Evaluation of the process of mainstreaming handicapped children into Project Head Start: Phase II final report. Silver Spring, Maryland: Applied Management Sciences, October 16, 1978.

Walker, D. K. Socioemotional measures for preschool and kindergarten children. San Francisco: Jossey-Bass, 1973.

Walker, D., Bane. M., & Bryk, A. The quality of the Head Start planned variation data. Cambridge, Massachusetts: Huron Institute, 1973.

Wright, B. D. Sample free test calibration and person measurement. Paper prepared for the 1967 Invitational Conference on Testing Problems, October 28, 1967.

Wright, B. D. & Stone, M. H. Best test design: Rasch measurement. Chicago: Mesa Press, 1979.

Zigler, E. America's Head Start program: An agenda for its second decade. Unpublished manuscript based upon a speech, 1977.

Zigler, E., & Valentine, J. Project Head Start: A legacy of the War on Poverty. New York: Free Press, 1979.

Zimiles, H. Assessing Head Start's impact on cognitive development of children. Research Division, Bank Street College of Education, New York, N.Y., 1977.